Universidade Federal de São Paulo

Instituto de Ciência e Tecnologia

Mestrado em Matemática Pura e Aplicada

IMAGENS DE POLINÔMIOS MULTILINEARES SOBRE ALGUMAS SUBÁLGEBRAS DE MATRIZES

Pedro Souza Fagundes

São José dos Campos

2019

Pedro Souza Fagundes

Imagens de polinômios multilineares sobre algumas subálgebras de matrizes

Dissertação apresentada à Universidade Federal de São Paulo - Instituto de Ciência e Tecnologia como parte dos requisitos para a obtenção do título de Mestre em Matemática Pura e Aplicada.
Orientador: Thiago Castilho de Mello
Coorientador: Mikhail Aleksandrovich Chebotar

São José dos Campos

2019

Fagundes, Pedro Souza

Imagens de polinômios multilineares sobre algumas subálgebras de matrizes / Pedro Souza Fagundes. - São José dos Campos, 2019.

xi, 86f.

Dissertação (Mestre) - Universidade Federal de São Paulo, Instituto de Ciência e Tecnologia. Programa de Pós-Graduação em Matemática Pura e Aplicada.

Titulo em inglês: The images of multilinear polynomials on some matrix subalgebras.

1. Conjectura de Lvov-Kaplansky. 2. polinômios multilineares. 3. matrizes triangulares superiores. 4. matrizes estritamente triangulares superiores.

Universidade Federal de São Paulo

Instituto de Ciência e Tecnologia

Programa de Pós-Graduação em Matemática Pura e Aplicada

Chefe do Departamento: Prof. Dr. Eduardo Antonelli

Coordenador do Programa: Prof. Dr. Thiago Castilho de Mello

Apoio Financeiro: FAPESP e CAPES

Pedro Souza Fagundes

IMAGENS DE POLINÔMIOS MULTILINEARES SOBRE ALGUMAS SUBÁLGEBRAS DE MATRIZES

Presidente da banca:

Prof. Dr. Thiago Castilho de Mello

Thiago Castilho de Mello

Banca Examinadora:

Prof. Dr. Angelo Calil Bianchi

Angelo

Prof. Dr. Dimas José Gonçalves

Dimas José Gonçalves

Prof. Dr. Plamen Emilov Kochloukov

À minha família.

Agradecimentos

Sou grato à minha família por me incentivar aos estudos, por acreditar em minhas decisões e me conceder todo um suporte que me manteve em pé nos momentos difíceis.

Sou grato aos meus amigos, antigos e novos, que me fizeram rir muito e deixaram a trajetória do mestrado mais feliz.

Sou grato ao professor Thiago de Mello por me orientar, e por ter sido tão prestativo, atencioso e paciente. Também agradeço ao professor Mikhail Chebotar por ter compartilhado comigo a sua infinita experiência. Ambos foram como pais para mim.

Sou grato a todos os professores do ICT-UNIFESP que contribuíram em minha formação matemática durante o mestrado.

Sou grato aos professores Angelo Bianchi, Dimas Gonçalves e Plamen Kochloukov por aceitarem o pedido de comporem a banca e por terem sugerido valiosas correções que certamente melhoraram o texto.

Sou grato ao apoio financeiro concedido pela FAPESP sob os processos nº 2016/09496-7 e nº 2017/16864-5, bem como à CAPES.

"Go, go, go, figure out, figure out, but don't stop moving."
David Guetta and Sia

Resumo

Nesta dissertação iremos estudar imagens de polinômios multilineares sobre a álgebra das matrizes. Apresentaremos resultados de Shoda, Albert e Muckenhoupt, Mesyan e Buzinski e Winstanley para estudarmos os casos em que os polinômios multilineares possuem grau até quatro. Também descreveremos as imagens de polinômios multilineares de grau até quatro sobre a álgebra das matrizes triangulares superiores bem como as imagens de polinômios multilineares de grau arbitrário sobre a álgebra das matrizes estritamente triangulares superiores. Além disso, estudaremos alguns resultados de Brešar e Klep sobre a relação entre o subespaço gerado pela imagem de polinômios não comutativos em álgebras e ideais de Lie.

Palavras-chave: Conjectura de Lvov-Kaplansky, polinômios multilineares, matrizes triangulares superiores, matrizes estritamente triangulares superiores.

Abstract

In this dissertation we will study the images of multilinear polynomials on matrix algebras. We will present results from Shoda, Albert and Muckenhoupt, Mesyan and Buzinski and Winstanley to study the cases where the multilinear polynomials have degree up to four. We will also describe the images of multilinear polynomials of degree up to four on the upper triangular matrix algebra as well as the images of multilinear polynomials of arbitrary degree on the strictly upper triangular matrix algebra. Moreover, we will study some results from Brešar and Klep about the relation between the linear span of the images of noncommutative polynomials on algebras and Lie ideals.

Keywords: Lvov-Kaplansky conjecture, multilinear polynomials, upper triangular matrices, strictly upper triangular matrices.

Sumário

INTRODUÇÃO

Dados um polinômio $f(x_1, \ldots, x_m)$ em variáveis não comutativas com coeficientes sobre um corpo $\mathbb{K}$ e uma $\mathbb{K}$-álgebra $\mathcal{A}$, definimos a imagem de f sobre $\mathcal{A}$ como a imagem da função

$$\begin{array}{rccc} \tilde{f}: & \mathcal{A}^m & \rightarrow & \mathcal{A} \\ & (a_1, \ldots, a_m) & \mapsto & f(a_1, \ldots, a_m) \end{array}$$

e a denotaremos por $f(\mathcal{A})$. Estudar a imagem de um polinômio é de grande relevância em álgebra não comutativa. Por exemplo, identidades polinomiais e polinômios centrais podem ser definidos em termos de imagens de polinômios.

Quando consideramos $\mathcal{A}$ como sendo a álgebra das matrizes sobre corpos finitos, então já entendemos bem o que ocorre com as imagens de polinômios não comutativos sobre $\mathcal{A}$. Em 1990, Chuang [8] provou que um subconjunto formado por tais matrizes é a imagem de um polinômio sem termo constante se e, somente se, esse subconjunto contém a matriz nula e é fechado sob conjugação.

Para o caso em que o polinômio f seja multilinear e a álgebra $\mathcal{A}$ seja a álgebra das matrizes $M_n(\mathbb{K})$, Lvov [13] conjecturou que $f(M_n(\mathbb{K}))$ é um espaço vetorial. Como veremos no Capítulo 2, essa conjectura, também atribuída à Kaplansky, pode ser reformulada equivalentemente como: $f(M_n(\mathbb{K}))$ é $\{0\}, \mathbb{K}, sl_n(\mathbb{K})$ ou $M_n(\mathbb{K})$, em que $\mathbb{K}$ denota o espaço das matrizes escalares e $sl_n(\mathbb{K})$ o espaço das matrizes de traço zero.

Um dos mais antigos resultados relacionados à Conjectura de Lvov-Kaplansky foi dado em 1936, quando Shoda [24] descreveu a imagem do polinômio $f(x_1, x_2) = x_1x_2 - x_2x_1$ sobre matrizes com entradas num corpo de característica zero, provando que toda matriz de traço zero pode ser escrita como um comutador. Em 1957, Albert e Muckenhoupt [1] generalizaram o resultado de Shoda para corpos quaisquer. Tal descrição é o passo de maior relevância para a prova da veracidade da conjectura para polinômios multilineares de grau dois. Vale notar que tal resultado não é verdade quando consideramos a álgebra das matrizes com entradas num anel qualquer. De fato, Rosset e Rosset [23] forneceram uma matriz em $M_2(R)$, em que R é um anel de polinômios, que não pode ser escrita como um comutador entre duas matrizes em $M_2(R)$.

O maior progresso na Conjectura de Lvov-Kaplansky foi feito em 2012 por Kanel-

Belov, Malev e Rowen [16], quando eles a estabeleceram para matrizes de ordem dois com entradas em corpos quadraticamente fechados. Ainda sobre matrizes de ordem dois, Malev [21] estabeleceu a conjectura sobre o corpo dos números reais e ainda provou que, sobre corpos quaisquer, a imagem de polinômios multilineares será zero, o centro ou conterá as matrizes de traço zero. Em 2016, esses mesmos três autores (ver [17]) estudaram a conjectura sobre $M_3(\mathbb{K})$, em que $\mathbb{K}$ é um corpo algebricamente fechado, e obtiveram, além dos quatro subespaços mencionados no 2º parágrafo, outros conjuntos como candidatos à imagem de polinômios multilineares, como, por exemplo, conjuntos densos em $M_3(\mathbb{K})$. Porém, não foi apresentado nenhum polinômio cuja imagem seja um desses novos conjuntos, o que deixa a Conjectura de Lvov-Kaplansky ainda um problema em aberto, mesmo para $M_3(\mathbb{K})$.

Recentemente, problemas relacionados à descrição de imagens de polinômios multilineares sobre a álgebra das matrizes foram bastante estudados. Em 2013, Mesyan [22] provou que a imagem de polinômios multilineares não nulos de grau três sobre a álgebra das matrizes $M_n(\mathbb{K})$, com algumas poucas condições sobre $\mathbb{K}$, sempre contém as matrizes de traço zero, e, ainda nesse ano, Buzinski e Winstanley [7] obtiveram um resultado similar para polinômios multilineares não nulos de grau quatro sobre $M_n(\mathbb{K})$, em que $n \geqslant 3$ e $\mathbb{K}$ um corpo algebricamente fechado de característica zero. Ainda sobre polinômios multilineares de grau três, Dykema e Klep [10] decreveram computacionalmente as imagens de tais polinômios (com coeficientes em $\mathbb{C}$) sobre matrizes $n \times n$ (com entradas em $\mathbb{C}$), em que n é par ou menor que 17.

Algumas variações da Conjectura de Lvov-Kaplansky também têm sido exploradas. Citamos, por exemplo, as descrições das imagens de polinômios de Lie multilineares de graus três e quatro sobre algumas álgebras de Lie clássicas [3], das imagens de polinômios de Lie multilineares de grau até quatro sobre a álgebra das matrizes [25], das imagens de polinômios de Lie homogêneos sobre matrizes de ordem 2 [18] e das imagens de polinômios de Jordan multilineares de grau três sobre a álgebra de Jordan das matrizes simétricas [20].

O objetivo principal desta dissertação é apresentar alguns resultados que respondem positivamente ou fazem um importante progresso na Conjectura de Lvov-Kaplansky, bem como descrever as imagens de polinômios multilineares de grau até quatro sobre a álgebra das matrizes triangulares superiores e as imagens de polinômios multilineares de grau arbitrário sobre a álgebra das matrizes estritamente triangulares superiores.

A dissertação está estruturada da seguinte forma: no Capítulo 1 são apresentados os principais conceitos e resultados preliminares sobre anéis e álgebras, bem como estabelecidas as notações que serão usadas nos demais capítulos. O Capítulo 2 terá como principal referência o artigo [6] do Brĕsar e Klep. Provaremos que o subespaço gerado pela imagem de um polinômio não comutativo sobre uma álgebra é um ideal de Lie, e apresentaremos uma classificação dos ideais de Lie das álgebras associativas centrais simples de dimensão finita. Partindo dessa classificação, descreveremos quais polinômios possuem o subespaço gerado pela imagem igual a um ideal de Lie dado.

O Capítulo 3 será dedicado à descrição das imagens dos polinômios multilineares de grau 2 sobre a álgebra das matrizes, bem como a um estudo completo dos artigos [22] e [7] que provam que as matrizes de traço zero pertencem às imagens de polinômios multilineares de graus 3 e 4 sobre a álgebra das matrizes de ordem n, em que $n \geqslant 2$ e $n \geqslant 3$, respectivamente. Além disso, iremos apresentar uma correção original de um erro em um lema do artigo [7].

Os próximos capítulos são dedicados à apresentação de resultados originais relacionados à Conjectura de Lvov-Kaplansky. No Capítulo 4 conjecturamos que a imagem de um polinômio multilinear sobre a álgebra das matrizes triangulares superiores é um espaço vetorial e provaremos essa conjectura para polinômios de grau até 4. Os resultados deste capítulo resultaram em um artigo, aceito para publicação no periódico Operators and Matrices (ver [12]). No Capítulo 5, iremos descrever as imagens de polinômios multilineares de grau arbitrário sobre a álgebra das matrizes estritamente triangulares superiores e tal resultado encontra-se publicado no periódico Linear Algebra and its Applications (ver [11]).

Capítulo 1

CONCEITOS PRELIMINARES

O principal objetivo deste capítulo será apresentar as definições, propriedades e alguns resultados básicos em álgebra, que serão utilizados no decorrer do texto. As principais referências são os livros [4], [5], [9] e [19].

1.1 Anéis

Definição 1.1. *Um conjunto não vazio* $\mathcal{R}$ *munido de duas operações binárias* $+ : \mathcal{R} \times \mathcal{R} \to \mathcal{R}$ *e* $\cdot : \mathcal{R} \times \mathcal{R} \to \mathcal{R}$ *é chamado de* **anel** *se valem as seguintes propriedades:*

1. $(a + b) + c = a + (b + c), \forall a, b, c \in \mathcal{R}$;

2. *existe* $0 \in \mathcal{R}$ *tal que* $a + 0 = 0 + a = a, \forall a \in \mathcal{R}$;

3. *para cada* $a \in \mathcal{R}$ *existe* $-a \in \mathcal{R}$ *tal que* $a + (-a) = (-a) + a = 0$;

4. $a + b = b + a, \forall a, b \in \mathcal{R}$;

5. $(a + b) \cdot c = a \cdot c + b \cdot c, \forall a, b, c \in \mathcal{R}$;

6. $a \cdot (b + c) = a \cdot b + a \cdot c, \forall a, b, c \in \mathcal{R}$.

O produto $a \cdot b$ será simplesmente denotado pela justaposição ab.

Definição 1.2. *Dizemos que um anel* $\mathcal{R}$*:*

1. *é* **associativo***, se* $(ab)c = a(bc)$*, para todos* $a, b, c \in \mathcal{R}$;

2. *é* **unitário***, se existe* $1 \in \mathcal{R}$ *tal que* $1a = a1 = a$*, para todo* $a \in \mathcal{R}$;

3. *é* **de divisão***, se é unitário e para cada* $a \in \mathcal{R}$ *não nulo existir* $a^{-1} \in \mathcal{R}$ *tal que* $aa^{-1} = a^{-1}a = 1$;

4. *é* **comutativo***, se* $ab = ba$*, para todos* $a, b \in \mathcal{R}$*;*

5. *é um* **corpo***, se satisfizer as quatro propriedades acima.*

A menos que seja dito o contrário, estaremos supondo durante o texto que os anéis são associativos.

Exemplo 1.3. *Um anel que será fortemente explorado nesta dissertação é o anel das matrizes* $M_n(\mathbb{K})$*, em que* $\mathbb{K}$ *é um corpo qualquer. Observamos que tal anel é associativo com unidade, porém não comutativo para* $n > 1$*. Destacamos também outros dois anéis, a saber: o anel das matrizes triangulares superiores* UT_n *(matrizes cujas entradas* (i, j) *são nulas para* $i > j$*) e o anel das matrizes estritamente triangulares superiores* $UT_n^{(0)}$ *(matrizes cujas entradas* (i, j) *são nulas para* $i \geqslant j$*).*

Definição 1.4. *Um corpo é dito ser* **algebricamente fechado** *se todo polinômio comutativo não constante em uma variável sobre esse corpo possui raiz.*

Definição 1.5. *Dado um corpo* $\mathbb{K}$*, o fecho algébrico de* $\mathbb{K}$ *é um corpo* $\overline{\mathbb{K}}$ *que satisfaz:*

(i) $\overline{\mathbb{K}}$ *é algebricamente fechado;*

(ii) existe um homomorfismo injetivo $\phi : \mathbb{K} \to \overline{\mathbb{K}}$*, em que todo elemento de* $\overline{\mathbb{K}}$ *é raiz de um polinômio com coeficientes em* $\phi(\mathbb{K})$*.*

Exemplo 1.6. *O corpo dos números complexos* $\mathbb{C}$ *é algebricamente fechado e é o fecho algébrico do corpo dos números reais* $\mathbb{R}$*.*

Definição 1.7. *O subconjunto* $Z(\mathcal{R}) = \{a \in \mathcal{R} | ab = ba, \forall b \in \mathcal{R}\}$ *do anel* $\mathcal{R}$ *é chamado o* **centro** *de* $\mathcal{R}$*.*

Exemplo 1.8. *Se* $\mathcal{R}$ *é um anel comutativo, então* $Z(\mathcal{R}) = \mathcal{R}$*.*

O próximo exemplo, que será utilizado futuramente, descreve o centro do anel das matrizes sobre um anel com unidade.

Exemplo 1.9. *Seja* $\mathcal{R}$ *um anel com unidade. Então* $Z(M_n(\mathcal{R})) = \{\lambda I_n | \lambda \in Z(\mathcal{R})\}$*, em que* I_n *denota a matriz identidade de ordem* n*.*

Durante a justificativa deste exemplo (e também em todo o texto), iremos denotar as matrizes unitárias por $e_{i,j}$*, isto é, matrizes com entrada* (i, j) *igual à 1 e zero nas demais. É fácil ver que* $\{\lambda I_n | \lambda \in Z(\mathcal{R})\} \subset Z(M_n(\mathcal{R}))$*. Para a inclusão contrária, seja* $A = (a_{i,j}) \in Z(M_n(\mathcal{R}))$*. Se em* A *existisse uma entrada* $a_{i,j} \neq 0$ *com* $i < j$*, então considerando a matriz* $e_{j,j} \in M_n(\mathcal{R})$ *temos que* $Ae_{j,j} \neq e_{j,j}A$*, pois na entrada* (i, j) *de* $Ae_{j,j}$ *temos* $a_{i,j} \neq 0$ *enquanto esta mesma entrada na matriz* $e_{j,j}A$ *é nula. Portanto, devemos* *ter $a_{i,j} = 0$ para $i < j$. Analogamente, $a_{i,j} = 0$ para $i > j$. Logo, A é uma matriz diagonal. Se em A tivéssemos $a_{i,i} \neq a_{j,j}$ para $i \neq j$, então*

$$Ae_{i,j} = \sum_{k=1}^{n} a_{k,k}e_{k,k}e_{i,j} = a_{i,i}e_{i,j} \neq a_{j,j}e_{i,j} = \sum_{k=1}^{n} a_{k,k}e_{i,j}e_{k,k} = e_{i,j}A.$$

Isto nos permite concluir que $A = \lambda I_n$ para algum $\lambda \in \mathcal{R}$. Agora, dado $\eta \in \mathcal{R}$, considerando a matriz ηI_n e que $A \in Z(M_n(\mathcal{R}))$, devemos ter $(\lambda I_n)(\eta I_n) = (\eta I_n)(\lambda I_n)$, isto é, $\lambda\eta = \eta\lambda$. Provamos assim $\lambda \in Z(\mathcal{R})$.

Definição 1.10. *Seja $\mathcal{R}$ um anel e $\mathcal{R}_1$ um subconjunto não vazio de $\mathcal{R}$.*

1. *$\mathcal{R}_1$ é um* **subanel** *de $\mathcal{R}$ se $\mathcal{R}_1$ é um anel cujas operações são as de $\mathcal{R}$ restritas a $\mathcal{R}_1$;*
2. *$\mathcal{R}_1$ é um* **ideal à esquerda** *de $\mathcal{R}$ se $a - b \in \mathcal{R}_1$ para todos $a, b \in \mathcal{R}_1$ e $ab \in \mathcal{R}_1$ para todos $a \in \mathcal{R}, b \in \mathcal{R}_1$;*
3. *$\mathcal{R}_1$ é um* **ideal à direita** *de $\mathcal{R}$ se $a - b \in \mathcal{R}_1$ para todos $a, b \in \mathcal{R}_1$ e $ba \in \mathcal{R}_1$ para todos $a \in \mathcal{R}, b \in \mathcal{R}_1$;*
4. *$\mathcal{R}_1$ é um* **ideal bilateral** *(ou simplesmente* **ideal***) se $\mathcal{R}_1$ é um ideal à esquerda e à direita de $\mathcal{R}$.*

Exemplo 1.11.

1. *$\{0\}$ e $\mathcal{R}$ são subanéis e ideais triviais do anel $\mathcal{R}$;*
2. *o anel $\left\{ \begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix} | a, b \in \mathbb{K} \right\}$ possui o subanel $\left\{ \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} | a \in \mathbb{K} \right\}$ como um ideal à direita;*
3. *o anel $\left\{ \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} | a, b \in \mathbb{K} \right\}$ possui o subanel $\left\{ \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} | a \in \mathbb{K} \right\}$ como um ideal à esquerda;*
4. *$UT_n^{(0)}$ é um ideal de UT_n.*

Sejam D um anel de divisão e I um ideal não nulo de D. Portanto, existe $a \in I - \{0\}$ e assim $1 = aa^{-1} \in I$. Desde que todo elemento $b \in D$ pode ser escrito como $b = b \cdot 1$, então obtemos $I = D$. Isto prova que D possui apenas ideais triviais e nos motiva a seguinte definição.

Definição 1.12. *Um anel $\mathcal{R}$ é dito* **simples** *se $\mathcal{R}^2 \neq 0$ e os únicos ideais de $\mathcal{R}$ são os triviais.*

Na definição acima, $\mathcal{R}^2 \neq 0$ é uma condição técnica e significa que existem $x, y \in \mathcal{R}$ tais que $xy \neq 0$.

Como vimos logo acima, concluímos que todo anel de divisão é simples. Vejamos um outro exemplo.

Exemplo 1.13. *O anel $M_n(D)$ é simples, quando D é um anel de divisão. De fato, seja I um ideal não nulo de $M_n(D)$. Portanto existe uma matriz não nula $A = (a_{i,j}) \in I$. Seja $a_{k,l} \neq 0$. Portanto, $a_{k,l}e_{i,j} = e_{i,k}Ae_{l,j} \in I$, quaisquer que sejam $i, j \in \{1, \ldots, n\}$. Segue que $de_{i,j} = (da_{k,l}^{-1})e_{i,i}a_{k,l}e_{i,j}$ pertence à I, qualquer que seja $d \in D$, o que prova $I = M_n(D)$.*

Estabelecemos agora algumas notações para certos ideais de um anel $\mathcal{R}$:

1. $\mathcal{R}a = \{ra | r \in \mathcal{R}\}$ é um ideal à esquerda de $\mathcal{R}$;
2. $a\mathcal{R} = \{ar | r \in \mathcal{R}\}$ é um ideal à direita de $\mathcal{R}$;
3. $\mathcal{R}a\mathcal{R} = \{$somas finitas de elementos da forma $r_1ar_2 | r_1, r_2 \in \mathcal{R}\}$ é um ideal de $\mathcal{R}$.
4. se I e J são ideais de $\mathcal{R}$, então IJ é o ideal formado pelas somas finitas de elementos da forma uv para $u \in I$ e $v \in I$.

Lema 1.14. *Seja $\mathcal{R}$ um anel. Então as seguintes condições são equivalentes:*

(i) Para todos $a, b \in \mathcal{R}$ tais que $a\mathcal{R}b = 0$, temos $a = 0$ ou $b = 0$;

(ii) Para todos ideais à esquerda I e J de $\mathcal{R}$ tais que $IJ = 0$, temos $I = 0$ ou $J = 0$;

(iii) Para todos ideais à direita I e J de $\mathcal{R}$ tais que $IJ = 0$, temos $I = 0$ ou $J = 0$;

(iv) Para todos ideais I e J de $\mathcal{R}$ tais que $IJ = 0$, temos $I = 0$ ou $J = 0$.

Demonstração. Provemos inicialmente que $(i) \Rightarrow (ii)$. Sejam I e J ideais à esquerda de $\mathcal{R}$ tais que $IJ = 0$. Como $RJ \subset J$, então $IRJ = 0$. Se $I \neq 0$, sejam $a \in I - \{0\}$ e $b \in J$. Como $aRb = 0$, por (i) temos $b = 0$, ou seja, $J = 0$. A implicação $(i) \Rightarrow (iii)$ é análoga. As implicações $(ii) \Rightarrow (iv)$ e $(iii) \Rightarrow (iv)$ são triviais. Para concluir a prova, basta mostrarmos que $(iv) \Rightarrow (i)$. Sejam $a, b \in \mathcal{R}$ tais que $a\mathcal{R}b = 0$. Portanto, o produto dos ideais $\mathcal{R}a\mathcal{R}$ e $\mathcal{R}b\mathcal{R}$ é nulo, e pelo item (iv) obtemos que algum deles é nulo. Suponha, sem perda de generalidade, que $\mathcal{R}a\mathcal{R} = 0$. Segue que $\mathcal{R}a$ e $a\mathcal{R}$ são ideais que satisfazem $\mathcal{R}a \cdot \mathcal{R} = \mathcal{R} \cdot a\mathcal{R} = 0$. Pelo item (iv), $\mathcal{R}a = a\mathcal{R} = 0$. Logo, $\mathbb{Z} \cdot a$ é um ideal e $(\mathbb{Z} \cdot a)\mathcal{R} = 0$. Novamente pelo item (iv) obtemos $\mathbb{Z} \cdot a = 0$, ou seja, $a = 0$. $\square$

Definição 1.15. *Um anel é dito ser* **primo** *se satisfaz uma das condições do Lema 1.14.*

Observamos que todo anel simples é primo.

Definição 1.16. *Um elemento não nulo $a \in \mathcal{R}$ é dito ser* **regular** *se para todos $b, c \in \mathcal{R}$ tivermos: $ab = 0$ implica $b = 0$ e $ca = 0$ implica $c = 0$. Um anel que possui todos elementos não nulos regulares é chamado de* **domínio**.

Dado um domínio comutativo, é possível construir um corpo que o "contém" como subanel. Esse corpo é chamado de corpo de frações do domínio em questão. Tal construção pode ser encontrada na maioria dos textos base de álgebra. O nosso objetivo a partir de agora será construir o análogo do corpo de frações para um anel (possivelmente não comutativo) cujos elementos centrais não nulos sejam regulares.

Seja $\mathcal{R}$ um anel cujos elementos não nulos de seu centro Z sejam regulares.

Definimos em $\mathcal{R} \times (Z - \{0\})$ a seguinte relação:

$$(r_1, z_1) \sim (r_2, z_2) \Leftrightarrow r_1 z_2 = r_2 z_1.$$

É imediato checar que tal relação é de equivalência. Denotaremos por rz^{-1} a classe do elemento (r, z), e o conjunto das classes de equivalência módulo a relação acima será denotado por $Q_Z(\mathcal{R})$. Definimos nesse conjunto as seguintes operações (bem definidas):

$$\begin{aligned} r_1 z_1^{-1} + r_2 z_2^{-1} &:= (r_1 z_2 + r_2 z_1)(z_1 z_2)^{-1} \\ r_1 z_1^{-1} \cdot r_2 z_2^{-1} &:= r_1 r_2 (z_1 z_2)^{-1} \end{aligned}$$

Com as operações acima, $Q_Z(\mathcal{R})$ possui a estrutura de anel. Além disso, zz^{-1}, em que $z \in Z - \{0\}$, é o elemento neutro com respeito a multiplicação em $Q_Z(\mathcal{R})$. O centro de $Q_Z(\mathcal{R})$, que será denotado por $\hat{Z}$, consiste de elementos da forma zw^{-1}, em que $z, w \in Z$ e $w \neq 0$. Portanto, se $z, w \in Z - \{0\}$, então zw^{-1} possui inverso wz^{-1}, o que faz de $\hat{Z}$ um corpo. Observamos também que dado $z \in Z - \{0\}$, o mapa $r \mapsto (rz)z^{-1}$ é um homomorfismo injetor de $\mathcal{R}$ em $Q_Z(\mathcal{R})$.

Definição 1.17. *O anel $Q_Z(\mathcal{R})$ é chamado de* **anel dos quocientes centrais** *do anel $\mathcal{R}$.*

Exemplo 1.18. *Se $\mathcal{R}$ é um domínio comutativo, então $Q_Z(\mathcal{R})$ é o corpo de frações de $\mathcal{R}$. Prova-se também que $Q_Z(M_n(\mathbb{Z})) \simeq M_n(\mathbb{Q})$.*

1.2 Produtos tensoriais

O principal objetivo desta seção será definir e estudar algumas propriedades do produto tensorial entre dois espaços vetoriais. Para relembrar conceitos como os de espaços vetoriais e transformações lineares entre eles recomendamos [19].

Iremos definir o produto tensorial entre dois espaços vetoriais via construção. Portanto, iniciaremos com a seguinte definição de espaço vetorial livre.

Definição 1.19. *Seja X um conjunto não vazio e $\mathbb{K}$ um corpo. O $\mathbb{K}$-espaço vetorial livre sobre X, denotado por $\mathbb{K}(X)$, é definido como o conjunto das funções de X em $\mathbb{K}$ que assumem apenas uma quantidade finita de valores não nulos munido de uma soma e um produto por escalar usuais.*

Observamos que uma base para $\mathbb{K}(X)$ é o conjunto das funções f_x que vale 1 em $x \in X$ e zero nos demais elementos de X.

Sejam U e V dois $\mathbb{K}$-espaços vetoriais e considere $\mathcal{M} = \mathbb{K}(U \times V)$. Seja $\mathcal{M}_0$ o subespaço de $\mathcal{M}$ gerado pelos vetores da forma

$$\begin{gathered} f_{(u_1+u_2,v)} - f_{(u_1,v)} - f_{(u_2,v)}, \\ f_{(u,v_1+v_2)} - f_{(u,v_1)} - f_{(u,v_2)}, \\ f_{(\alpha u,v)} - \alpha f_{(u,v)}, \\ f_{(u,\alpha v)} - \alpha f_{(u,v)}, \end{gathered}$$

em que $u, u_1, u_2 \in U, v, v_1, v_2 \in V$ e $\alpha \in \mathbb{K}$.

Definição 1.20. *O produto tensorial $U \otimes V$ entre dois $\mathbb{K}$-espaços vetoriais U e V é definido como sendo o $\mathbb{K}$-espaço vetorial quociente $\mathcal{M}/\mathcal{M}_0$.*

Dados $u \in U$ e $v \in V$, denotamos $u \otimes v = f_{(u,v)} + \mathcal{M}_0$. Desde que as funções $f_{(u,v)}$ formam uma base para $\mathcal{M}$, então os vetores $u \otimes v$ formam um conjunto gerador para $U \otimes V$.

Proposição 1.21. *Sejam U, V e W três $\mathbb{K}$-espaços vetoriais. Se $U \otimes W = V \otimes W$, então $U = V$.*

Demonstração. Sejam $u \in U$ e $w \in W$. Então $u \otimes w \in V \otimes W$, o que implica em $u \otimes w = \sum_i \alpha_i v_i \otimes w_i$. Como $f_{(u,w)} \in u \otimes w$, segue pela última igualdade que $f_{(u,w)}$ é uma função com domínio em $V \times W$. Portanto devemos ter $U = V$. □

Encerramos esta seção com o seguinte lema sobre espaços vetoriais.

Lema 1.22. *Sejam V um $\mathbb{K}$-espaço vetorial e U um subespaço de V. Sejam $c_0, c_1, \ldots, c_n \in V$ tais que*

$$\sum_{i=0}^{n} \lambda^i c_i \in U$$

para pelo menos $n + 1$ escalares distintos λ. Então, $c_i \in U$ para todo i.

Demonstração. Sejam $\lambda_0, \lambda_1, \ldots, \lambda_n \in \mathbb{K}$ distintos entre si tais que $\sum_{i=0}^{n} \lambda_l^i c_i \in U$, para $l \in \{0, \ldots, n\}$. Logo, para cada $l \in \{0, \ldots, n\}$,

$$\sum_{i=0}^{n} \lambda_l^i \bar{c}_i = 0 \quad \text{em} \quad V/U.$$

As equações acima podem ser representadas matricialmente como

$$\begin{pmatrix} 1 & \lambda_0 & \cdots & \lambda_0^n \\ 1 & \lambda_1 & \cdots & \lambda_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^n \end{pmatrix} \begin{pmatrix} \bar{c}_0 \\ \bar{c}_1 \\ \vdots \\ \bar{c}_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Desde que os escalares $\lambda_l, l \in \{0, \ldots, n\}$, são distintos entre si, então a matriz de Vandermonde acima possui determinante não nulo. Portanto, $\bar{c}_i = 0$ em V/U, o que prova o lema. □

1.3 Álgebras

Definição 1.23. *Seja $\mathbb{K}$ um corpo. Uma $\mathbb{K}$-álgebra $\mathcal{A}$ (ou simplesmente álgebra) é um $\mathbb{K}$-espaço vetorial munido de uma operação binária $\cdot : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ (chamada produto) e tal que para quaisquer $a, b, c \in \mathcal{A}$ e $\alpha \in \mathbb{K}$ valem:*

1. $(a + b) \cdot c = a \cdot c + b \cdot c$;
2. $a \cdot (b + c) = a \cdot b + a \cdot c$;
3. $\alpha(a \cdot b) = a \cdot (\alpha b) = (\alpha a) \cdot b$.

O produto $a \cdot b$ será simplesmente denotado por ab.

Definição 1.24. *Dizemos que uma álgebra $\mathcal{A}$*

1. *é associativa, se $(ab)c = a(bc)$, para todos $a, b, c \in \mathcal{A}$;*
2. *é comutativa, se $ab = ba$, para todos $a, b \in \mathcal{A}$;*
3. *é unitária, se existe $1 \in \mathcal{A}$ tal que $1a = a1 = a$, para todo $a \in \mathcal{A}$.*

A menos que seja dito o contrário, as álgebras consideradas durante o texto serão associativas.

Exemplo 1.25. *Se $n > 1$, então o anel das matrizes $M_n(\mathbb{K})$ é uma $\mathbb{K}$-álgebra associativa, não comutativa e unitária.*

Definição 1.26. *Seja $\mathfrak{C}$ uma classe de álgebras e $\mathcal{A} \in \mathfrak{C}$ uma álgebra gerada (como álgebra) por um conjunto X. $\mathcal{A}$ é dita ser* **livre** *na classe $\mathfrak{C}$, livremente gerada por X, se para qualquer álgebra $\mathcal{B} \in \mathfrak{C}$ e qualquer função $\varphi_0 : X \to \mathcal{B}$, existe um único homomorfismo $\varphi : \mathcal{A} \to \mathcal{B}$ que estende φ_0.*

Sejam $\mathcal{A}$ e $\mathcal{B}$ duas álgebras livres em uma classe $\mathcal{C}$, em que $\mathcal{A}$ é livremente gerada por um conjunto X e $\mathcal{B}$ é livremente gerada por um conjunto Y. Se X e Y possuem mesma cardinalidade, então $\mathcal{A}$ e $\mathcal{B}$ são isomorfas. De fato, seja $\phi : X \to Y$ uma bijeção, e considere as funções $X \to \mathcal{B}$ e $Y \to \mathcal{A}$ dadas por $x \mapsto \phi(x)$ e $y \mapsto \phi^{-1}(y)$, respectivamente. Portanto tais funções podem ser estendidas para homomorfismos $\varphi_1 : \mathcal{A} \to \mathcal{B}$ e $\varphi_2 : \mathcal{B} \to \mathcal{A}$. Então, $\varphi_2 \circ \varphi_1 : \mathcal{A} \to \mathcal{A}$ é um homomorfismo que satisfaz $\varphi_2 \circ \varphi_1(x) = x$, para todo $x \in X$, e portanto $\varphi_2 \circ \varphi_1 = id_{\mathcal{A}}$. Analogamente prova-se que $\varphi_1 \circ \varphi_2 = id_{\mathcal{B}}$, o que nos permite concluir que φ_1 é um isomorfismo. Portanto, uma álgebra livre, livremente gerada por um conjunto X, é única a menos de isomorfismo. A recíproca desta observação também vale, ou seja, se $\mathcal{A}$ e $\mathcal{B}$, como acima, são isomorfas, então X e Y possuem a mesma cardinalidade. Isto será provado após o seguinte exemplo.

Exemplo 1.27. *Seja $X = \{x_1, x_2, \dots\}$ um conjunto enumerável de variáveis não comutativas. O $\mathbb{K}$-espaço vetorial gerado pelos monômios da forma $x_{i_1} \cdots x_{i_n}, x_{i_j} \in X, n > 0$, e multiplicação dada por justaposição*

$$(x_{i_1} \cdots x_{i_k})(x_{i_{k+1}} \cdots x_{i_n}) = x_{i_1} \cdots x_{i_n}$$

é uma álgebra livre na classe das álgebras associativas com unidade, em que a unidade é o monômio formado por zero variáveis. Essa álgebra é denotada por $\mathbb{K}\langle X\rangle$ e chamada de **álgebra associativa livre**. *Os elementos de $\mathbb{K}\langle X\rangle$ são chamados de polinômios não comutativos, ou simplesmente polinômios, se o contexto for claro.*

No exemplo acima, se X for formado por variáveis comutativas, então obtemos uma álgebra livre na classe das álgebras associativas, comutativas e com unidade, denotada por $\mathbb{K}[X]$.

Como prometido, provemos a recíproca da afirmação acima. Pelo Exemplo 1.27, a menos de isomorfismo, podemos supor que $\mathcal{A} = \mathbb{K}\langle X\rangle$ e $\mathcal{B} = \mathbb{K}\langle Y\rangle$. Definimos os seguintes subespaços de $\mathbb{K}\langle X\rangle$: para cada $k \in \mathbb{N}$, $\omega^k(\mathbb{K}\langle X\rangle)$ será o subespaço gerado por todos os monômios de grau pelo menos k. Uma primeira observação sobre tais subespaços é $\omega^{k+1}(\mathbb{K}\langle X\rangle) \subset \omega^k(\mathbb{K}\langle X\rangle)$, para todo $k \in \mathbb{N}$. Analogamente definimos $\omega^k(\mathbb{K}\langle Y\rangle)$. Por hipótese, seja $\phi : \mathbb{K}\langle X\rangle \to \mathbb{K}\langle Y\rangle$ um isomorfismo. Para cada $x_i \in X$, seja $\phi(x_i) = f_i(y_{n_1} \dots, y_{n_i})$. Considere o isomorfismo $\theta : \mathbb{K}\langle X\rangle \to \mathbb{K}\langle X\rangle$, definido por $\theta(x_i) = x_i - f_i(0, \dots, 0), x_i \in X$. Portanto, a composição $\psi = \phi \circ \theta$ também é um isomorfismo, e satisfaz $\psi(x_i) = f_i(y_{n_1}, \dots, y_{n_i}) - f_i(0, \dots, 0)$. Provemos que $\psi(\omega^k(\mathbb{K}\langle X\rangle)) = \omega^k(\mathbb{K}\langle Y\rangle)$, para todo $k \in \mathbb{N}$. Iniciamos com $k = 1$. É fácil ver que $\psi(\omega^1(\mathbb{K}\langle X\rangle)) \subset \omega^1(\mathbb{K}\langle Y\rangle)$. Para a inclusão contrária, note que qualquer polinômio em $\omega^1(\mathbb{K}\langle Y\rangle)$ não pode ser realizado como pré-imagem por ψ de algum polinômio com termo constante não nulo em $\mathbb{K}\langle X\rangle$. Suponha agora $k > 1$. Iremos provar que $\psi(\omega^k(\mathbb{K}\langle X\rangle)) \supset \omega^k(\mathbb{K}\langle Y\rangle)$, já que a outra inclusão é óbvia. Portanto, seja $g \in \omega^k(\mathbb{K}\langle Y\rangle)$.

Sabemos que existe $f \in \mathbb{K}\langle X\rangle$ tal que $\psi(f) = g$. Queremos mostrar que $f \in \omega^k(\mathbb{K}\langle X\rangle)$. Para tal fim, consideremos os dois itens abaixo:

1. Para cada $x_i \in X$, temos $\psi(x_i) \in \omega^1(\mathbb{K}\langle Y\rangle)\backslash\omega^2(\mathbb{K}\langle Y\rangle)$.

Obviamente $\psi(x_i) = f_i(y_{n_1}, \ldots, y_{n_i}) - f_i(0, \ldots, 0) \in \omega^1(\mathbb{K}\langle Y\rangle)$. Desde que $\psi(x_i) = \psi(f_i(\psi^{-1}(y_{n_1}), \ldots, \psi^{-1}(y_{n_i})) - f_i(0, \ldots, 0))$, segue que $x_i = f_i(\psi^{-1}(y_{n_1}), \ldots, \psi^{-1}(y_{n_i})) - f_i(0, \ldots, 0)$. Se ocorresse $\psi(x_i) \in \omega^2(\mathbb{K}\langle Y\rangle)$, então pela equação anterior teríamos $x_i \in \omega^2(\mathbb{K}\langle X\rangle)$, o que é um absurdo. Isso conclui o item 1.

2. Se $f(x_1, \ldots, x_m) \in \mathbb{K}\langle X\rangle$ e $h_1, \ldots, h_m \in \omega^1(\mathbb{K}\langle Y\rangle)\backslash\omega^2(\mathbb{K}\langle Y\rangle)$, então $f(x_1, \ldots, x_m) \in \omega^k(\mathbb{K}\langle X\rangle)\backslash\omega^{k+1}(\mathbb{K}\langle X\rangle)$ se, e somente se, $f(h_1, \ldots, h_m) \in \omega^k(\mathbb{K}\langle Y\rangle)\backslash\omega^{k+1}(\mathbb{K}\langle Y\rangle)$.

Suponha $f(x_1, \ldots, x_m) \in \omega^k(\mathbb{K}\langle X\rangle)\backslash\omega^{k+1}(\mathbb{K}\langle X\rangle)$. Como $f(x_1, \ldots, x_m) \in \omega^k(\mathbb{K}\langle X\rangle)$ e $h_1, \ldots, h_m \in \omega^1(\mathbb{K}\langle Y\rangle)$, segue que $f(h_1, \ldots, h_m) \in \omega^k(\mathbb{K}\langle Y\rangle)$. Desde que $f(x_1, \ldots, x_m) \notin \omega^{k+1}(\mathbb{K}\langle X\rangle)$, então f possui um monômio de grau k e como $h_1, \ldots, h_m \notin \omega^2(\mathbb{K}\langle Y\rangle)$, então cada h_j deve conter algum monômio de grau 1. Portanto, $f(h_1, \ldots, h_m)$ possui algum monômio de grau k, ou seja, $f(h_1, \ldots, h_m) \notin \omega^{k+1}(\mathbb{K}\langle Y\rangle)$. Reciprocamente, note que se $f \notin \omega^k(\mathbb{K}\langle X\rangle)$, então f possui algum monômio de grau menor do que k, e como cada h_j possui algum monômio de grau 1, temos $f(h_1, \ldots, h_m) \notin \omega^k(\mathbb{K}\langle Y\rangle)$. Além disso, se $f(x_1, \ldots, x_m) \in \omega^{k+1}(\mathbb{K}\langle X\rangle)$, como $h_j \in \omega^1(\mathbb{K}\langle Y\rangle)$, segue que $f(h_1, \ldots, h_m) \in \omega^{k+1}(\mathbb{K}\langle Y\rangle)$. Isso conclui o item 2.

Portanto, como $f(\psi(x_1), \ldots, \psi(x_m)) = \psi(f(x_1, \ldots, x_m)) = g$, os dois lemas acima combinados nos forneçe $f \in \omega^k(\mathbb{K}\langle X\rangle)$. Concluímos agora que

$$\dim(\omega(\mathbb{K}\langle X\rangle)) = \dim(\omega(\mathbb{K}\langle Y\rangle)) \text{ e } \dim(\omega^2(\mathbb{K}\langle X\rangle)) = \dim(\omega^2(\mathbb{K}\langle Y\rangle))$$

e logo,

$$|X| = \dim(\omega(\mathbb{K}\langle X\rangle)/\omega^2(\mathbb{K}\langle X\rangle)) = \dim(\omega(\mathbb{K}\langle Y\rangle)/\omega^2(\mathbb{K}\langle Y\rangle)) = |Y|.$$

Isso conclui a justificativa de que se $\mathbb{K}\langle X\rangle$ e $\mathbb{K}\langle Y\rangle$ são isomorfas, então X e Y possuem a mesma cardinalidade.

Definiremos a seguir um operador linear em $\mathbb{K}\langle X\rangle$, que terá uma grande importância no estudo de imagens de polinômios. Dado um conjunto enumerável $X = \{x_1, x_2, \ldots\}$ de variáveis não comutativas, para cada $j \in \mathbb{N}$ definimos a seguinte função de X em $\mathbb{K}\langle X\rangle$

$$x_i \quad \mapsto \quad \delta_{ij} = \begin{cases} 1 & \text{se} \quad i = j \\ 0 & \text{se} \quad i \neq j \end{cases}$$

Podemos portanto definir uma função nos monômios de $\mathbb{K}\langle X\rangle$ da forma $x_{i_1}\cdots x_{i_n}$ pondo $x_{i_1}\cdots x_{i_k}\cdots x_{i_n} \mapsto \sum_{k=1}^{n} x_{i_1}\cdots \delta_{i_k j}\cdots x_{i_n}$. A extensão linear dessa função para $\mathbb{K}\langle X\rangle$, denotada por $\dfrac{\partial}{\partial x_j}$, é chamada de **derivada parcial** em relação à variável x_j.

O próximo resultado é de grande importância quando consideramos polinômios comutativos sobre corpos infinitos.

Lema 1.28. *Sejam $\mathbb{K}$ um corpo infinito e $p(y_1,\ldots,y_m) \in \mathbb{K}[y_1,\ldots,y_m]$. Se $p(\alpha_1,\ldots,\alpha_m) = 0$ para quaisquer $\alpha_1,\ldots,\alpha_m \in \mathbb{K}$, então $p(y_1,\ldots,y_m) = 0$.*

Demonstração. O resultado é certamente válido para polinômios em uma variável, já que tais polinômios possuem um número finito de raízes. Suponhamos então que o resultado seja válido para polinômios em $\mathbb{K}[y_1,\ldots,y_{m-1}]$. Seja $p(y_1,\ldots,y_m) \in \mathbb{K}[y_1,\ldots,y_m]$. Como $\mathbb{K}[y_1,\ldots,y_m] = (\mathbb{K}[y_1,\ldots,y_{m-1}])[y_m]$, então podemos escrever

$$p(y_1,\ldots,y_m) = p_0(y_1,\ldots,y_{m-1}) + p_1(y_1,\ldots,y_{m-1})y_m + \cdots + p_s(y_1,\ldots,y_{m-1})y_m^s,$$

em que $s \geqslant 1$ é um inteiro.

Fixado quaisquer $\alpha_1,\ldots,\alpha_{m-1} \in \mathbb{K}$ temos que $p(\alpha_1,\ldots,\alpha_{m-1},y_m) \in \mathbb{K}[y_m]$ e por hipótese, $p_i(\alpha_0,\ldots,\alpha_{m-1}) = 0$, para todo $i \in \{1,\ldots,s\}$. Portanto, a hipótese de indução nos garante que $p_i(y_1,\ldots,y_{m-1}) = 0$ para todo i, o que prova $p = 0$. $\square$

Definição 1.29. *(i) O grau de um monômio $m(x_1,\ldots,x_n) \in \mathbb{K}\langle X\rangle$ na variável x_i é o número de ocorrências dessa variável em m;*

(ii) O grau total de um monômio $m(x_1,\ldots,x_n) \in \mathbb{K}\langle X\rangle$ é a soma dos graus de m em cada variável x_i;

(iii) O grau de um polinômio $f(x_1,\ldots,x_n) \in \mathbb{K}\langle X\rangle$ é o maior grau total dentre todos os seus monômios;

(iv) Um polinômio $f(x_1,\ldots,x_n) \in \mathbb{K}\langle X\rangle$ é homogêneo de grau k_i na variável x_i se cada monômio de f possui grau k_i em x_i;

(v) Um polinômio $f(x_1,\ldots,x_n) \in \mathbb{K}\langle X\rangle$ é multi-homogêneo de multigrau $(k_1,\ldots,k_n)$ se f é homogêneo de grau k_i em x_i, para todo i;

(vi) Um polinômio $f(x_1,\ldots,x_n) \in \mathbb{K}\langle X\rangle$ é multilinear se f é multi-homogêneo de multigrau $(1,\ldots,1)$.

Definição 1.30. *(i) Uma $\mathbb{K}$-álgebra $\mathcal{A}$ é dita ser simples, se $\mathcal{A}$ é simples como anel;*

(ii) Uma $\mathbb{K}$-álgebra $\mathcal{A}$ é dita ser prima se $\mathcal{A}$ é prima como anel;

(iii) *A dimensão de uma $\mathbb{K}$-álgebra $\mathcal{A}$ é a dimensão de $\mathcal{A}$ como espaço vetorial;*

(iv) *Uma $\mathbb{K}$-álgebra $\mathcal{A}$ é dita ser central se o centro de $\mathcal{A}$ é isomorfo ao corpo $\mathbb{K}$.*

Destacamos o seguinte resultado sobre álgebras centrais simples de dimensão finita, cuja demonstração pode ser encontrada, por exemplo, em [5], p. 48.

Teorema 1.31. *Seja $\mathcal{A}$ uma $\mathbb{K}$-álgebra central simples de dimensão finita. Se $\overline{\mathbb{K}}$ é o fecho algébrico de $\mathbb{K}$, então $\mathcal{A} \otimes \overline{\mathbb{K}} \simeq M_n(\overline{\mathbb{K}})$, em que $n = \sqrt{\dim_{\mathbb{K}}(\mathcal{A})}$.*

Note que do isomorfismo acima, temos $\dim_{\mathbb{K}}(\mathcal{A}) = \dim_{\bar{\mathbb{K}}} \mathcal{A} \otimes \bar{\mathbb{K}} = n^2$, e assim segue a igualdade $n = \sqrt{\dim_{\mathbb{K}}(\mathcal{A})}$.

1.3.1 Álgebras com identidades polinomiais

Definição 1.32. *Um polinômio $f = f(x_1, \ldots, x_m) \in \mathbb{K}\langle X \rangle$ é dito ser uma* **identidade polinomial** *para uma $\mathbb{K}$-álgebra $\mathcal{A}$ se $f(a_1, \ldots, a_m) = 0$ para todo $a_1, \ldots, a_m \in \mathcal{A}$. Dizemos que uma $\mathbb{K}$-álgebra $\mathcal{A}$ é uma* **PI-álgebra** *se $\mathcal{A}$ satisfaz alguma identidade polinomial não nula. Um polinômio $f = f(x_1, \ldots, x_m) \in \mathbb{K}\langle X \rangle$ é dito ser* **central** *para uma $\mathbb{K}$-álgebra $\mathcal{A}$ se f não for uma identidade polinomial para $\mathcal{A}$ e $f(a_1, \ldots, a_m) \in Z(\mathcal{A})$, para quaisquer $a_1, \ldots, a_m \in \mathcal{A}$.*

Exemplo 1.33. *Seja S_n o grupo das permutações dos números naturais $1, \ldots, n$. Definimos o polinômio standard de grau n por $s_n(x_1, \ldots, x_n) = \sum_{\sigma \in S_n} (-1)^\sigma x_{\sigma(1)} \cdots x_{\sigma(n)}$, em que $(-1)^\sigma$ é o sinal da permutação σ. O famoso Teorema de Amitsur-Levitzki (ver [9], p. 79) nos garante que o polinômio standard s_{2n} é uma identidade polinomial para $M_n(\mathbb{K})$.*

Exemplo 1.34. *Levando em consideração que o comutador entre duas matrizes triangulares superiores de ordem n é uma matriz estritamente triangular superior e o produto de n matrizes estritamente triangulares superiores de ordem n é zero, segue que o polinômio $[x_1, x_2] \cdots [x_{2n-1}, x_{2n}]$ é uma identidade polinomial para UT_n.*

Exemplo 1.35. *Segue do Teorema de Cayley-Hamilton (ver [9], p. 89) que o polinômio $f(x_1, x_2) = [x_1, x_2]^2$ é central para $M_2(\mathbb{K})$.*

Observamos que se $f \in \mathbb{K}\langle X \rangle$ é uma identidade polinomial para uma $\mathbb{K}$-álgebra $\mathcal{A}$, então podemos obter através de f uma identidade polinomial multilinear para $\mathcal{A}$ de grau menor ou igual ao de f. O processo para obtenção de tal identidade é chamado de **processo de linearização** (ver [5], p. 148).

Proposição 1.36. *Todo polinômio $f \in \mathbb{K}\langle X \rangle$ de grau menor que $2n$ não é identidade polinomial de $M_n(\mathbb{K})$.*

Demonstração. O processo de linearização nos permite supor que f é multilinear. Suponha inicialmente que o grau de f seja $2n-1$. Portanto, podemos escrever

$$f(x_1,\ldots,x_n)=\sum_{\sigma\in S_{2n-1}}\alpha_\sigma x_{\sigma(1)}\cdots x_{\sigma(2n-1)},$$

em que $\alpha_\sigma\in\mathbb{K}$. Podemos supor, sem perda de generalidade, que $\alpha_{id}\neq 0$, em que id é a permutação identidade. Como o produto das matrizes $e_{1,1},e_{1,2},e_{2,2},e_{2,3},\ldots,e_{n-1,n},e_{n,n}$ é não nulo (e igual à $e_{1,n}$) se e somente se for realizado nessa ordem, segue que $f(e_{1,1},e_{1,2},\ldots,e_{n,n})=\alpha_{id}e_{1,n}\neq 0$. O caso em que o grau de f é menor que $2n-1$ pode ser tratado analogamente. $\square$

Dada uma álgebra $\mathcal{A}$, denotaremos por $T(\mathcal{A})$ o conjunto de todas as identidades polinomiais de $\mathcal{A}$. Observamos que $T(\mathcal{A})$ é um ideal de $\mathbb{K}\langle X\rangle$ e se $\varphi:\mathbb{K}\langle X\rangle\to\mathbb{K}\langle X\rangle$ é um homomorfismo, então $\varphi(T(\mathcal{A}))\subset T(\mathcal{A})$. Ideais de $\mathbb{K}\langle X\rangle$ que satisfazem essa propriedade são chamados de T-ideais.

O próximo lema é uma consequência imediata do Lema 1.22.

Lema 1.37. *Seja $\mathbb{K}$ um corpo infinito e $\mathcal{A}$ uma $\mathbb{K}$-álgebra. Se $f\in T(\mathcal{A})$, então o mesmo ocorre para suas componentes multi-homogêneas.*

Em [9], p. 23, é provado que se $\mathcal{A}$ é uma $\mathbb{K}$-álgebra associativa, então $\mathbb{K}\langle X\rangle/T(A)$ é a álgebra livre na classe de todas as álgebras associativas que satisfazem as mesmas identidades polinomiais de $\mathcal{A}$.

Sejam $n\geqslant 2$ um inteiro, $\mathbb{K}$ um corpo infinito e $\mathbb{K}[y_{p,q}^{(l)}|p,q\in\{1,\ldots,n\},l\in\mathbb{N}]$ uma álgebra de polinômios comutativos. Denotando

$$y_i=\begin{pmatrix} y_{1,1}^{(i)} & \cdots & y_{1,n}^{(i)}\\ \vdots & \ddots & \vdots\\ y_{n,1}^{(i)} & \cdots & y_{n,n}^{(i)}\end{pmatrix},$$

temos a seguinte definição:

Definição 1.38. *As matrizes $y_i\in M_n(\mathbb{K}[y_{p,q}^{(l)}])$ são chamadas de* **matrizes genéricas elementares***. A subálgebra $GM_n(\mathbb{K})$ de $M_n(\mathbb{K}[y_{p,q}^{(l)}])$ gerada por tais matrizes é chamada de* **álgebra das matrizes genéricas**.

Na próxima proposição iremos provar que a álgebra das matrizes genéricas é livre na classe de todas as álgebras que satisfazem as mesmas identidades polinomiais que a álgebra das matrizes.

Proposição 1.39. *Se $\mathbb{K}$ é um corpo infinito, então*

$$GM_n(\mathbb{K})\simeq\mathbb{K}\langle X\rangle/T(M_n(\mathbb{K})).$$

Demonstração. Considere o homomorfismo sobrejetor $\varphi : \mathbb{K}\langle X\rangle \to GM_n(\mathbb{K})$ tal que $\varphi(x_i) = y_i$. Provemos que $ker(\varphi) = T(M_n(\mathbb{K}))$. Seja $f(x_1, \ldots, x_m) \in ker(\varphi)$. Então $f(y_1, \ldots, y_m) = \varphi(f(x_1, \ldots, x_m)) = 0$. Sejam $A_i = (a_{p,q}^{(l)}) \in M_n(\mathbb{K}), i = 1, \ldots, m$, quaisquer matrizes. Sendo $\mathbb{K}[y_{p,q}^{(l)} | l \in \mathbb{N}]$ livre na classe das álgebras associativas comutativas unitárias, existe um homomorfismo $\psi : \mathbb{K}[y_{p,q}^{(l)} | l \in \mathbb{N}] \to \mathbb{K}$ tal que $\psi(y_{p,q}^{(l)}) = a_{p,q}^{(l)}$ para $l = 1, \ldots, m$ e $\psi(y_{p,q}^{(l)}) = 0$ para $l > m$. Portanto ψ induz um homomorfismo $\tilde{\psi} : GM_n(\mathbb{K}) \to M_n(\mathbb{K})$ em que $\tilde{\psi}(y_l) = A_l$, para $l \in \{1, \ldots, m\}$ e $\tilde{\psi}(y_l) = 0$ para $l > m$. Segue que

$$f(A_1, \ldots, A_m) = \tilde{\psi}(f(y_1, \ldots, y_m)) = \tilde{\psi}(0) = 0,$$

ou seja, $ker(\varphi) \subset T(M_n(\mathbb{K}))$. Reciprocamente, seja $f(x_1, \ldots, x_m) \in T(M_n(\mathbb{K}))$. Então para quaisquer matrizes $A_l, l = 1, \ldots, m$, temos que $0 = f(A_1, \ldots, A_m) = \tilde{\psi}(f(y_1, \ldots, y_m))$, ou seja, cada entrada de $f(y_1, \ldots, y_m)$ é um polinômio em variáveis comutativas que se anula sobre $\mathbb{K}$. Pelo Lema 1.28, cada um desses polinômios deve ser igual ao polinômio nulo, e portanto $f(y_1, \ldots, y_m) = 0$, o que prova que $f(x_1, \ldots, x_m) \in ker(\varphi)$. □

Encerramos esta seção com o conhecido Teorema de Posner, cuja demonstração pode ser encontrada em [5], p. 187.

Teorema 1.40 (Posner)**.** *Se $\mathcal{R}$ é um PI-anel primo, então $Q_Z(\mathcal{R})$ é uma álgebra central simples de dimensão finita sobre $\hat{Z}$.*

1.3.2 Álgebras de Lie

Definição 1.41. *Uma $\mathbb{K}$-álgebra $\mathcal{A}$ é chamada de Lie se satisfaz as seguintes condições:*

1. $aa = 0, \forall a \in \mathcal{A}$;
2. $(ab)c + (bc)a + (ca)b = 0, \forall a, b, c \in \mathcal{A}$.

O produto ab em uma álgebra de Lie será denotado por $[a, b]$.

Exemplo 1.42. *Se $\mathcal{A}$ é uma $\mathbb{K}$-álgebra associativa, então para $a, b \in \mathcal{A}$ definimos uma nova multiplicação em $\mathcal{A}$ dada por $[a, b] = ab - ba$ (o comutador entre a e b). A álgebra obtida é de Lie e a denotamos por $\mathcal{A}^{(-)}$.*

Denotaremos por $L(X)$ a subálgebra de Lie de $\mathbb{K}\langle X\rangle^{(-)}$ gerada por X. Pelo Teorema de Witt (ver [9], p. 14) temos que $L(X)$ é a álgebra de Lie livre, livremente gerada por X. Os elementos de X são chamados de monômios de grau 1. Se u é um monômio de grau r e v é um monômio de grau s, com u e v distintos, então $[u, v]$ é um monômio de grau $r + s$. Caso $u = v$, então $[u, v] = 0$.

Agora iremos definir indutivamente um conjunto ordenado R de monômios em $L(X)$. Ponha $X \subset R$ como sendo os monômios de grau 1 e defina uma ordem entre os elementos de X. Suponha definidos os monômios de grau $1, \ldots, n-1$ em R com uma ordem que satisfaça $u < v$ se o grau de u for menor que o grau de v. Sejam u e v monômios tais que $[u, v]$ possui grau n. Então $[u, v] \in R$, se:

(i) $u, v \in R$ e $u > v$;

(ii) se $u = [u_1, u_2]$, então $v \geqslant u_2$.

Basta então ordenar os monômios de grau n em R arbitrariamente.

O conjunto R definido acima é uma base de $L(X)$ e é conhecido como **base de Hall**. Uma demonstração de que tal conjunto é de fato uma base pode ser encontrada em [14].

Vejamos a seguir dois exemplos de bases de Hall que serão usadas durante o texto.

Exemplo 1.43. *Seja $X = \{x_1, x_2, x_3\}$. Construiremos uma base de Hall dos polinômios de Lie multilineares em $\mathbb{K}\langle X\rangle$ de grau igual a 3.*

Ordenamos X pondo $x_1 < x_2 < x_3$. Portanto, os polinômios de Lie multilineares de grau 2 são: $[x_2, x_1], [x_3, x_1]$ e $[x_3, x_1]$. Segue que os polinômios de Lie multilineares de grau 3 são: $[[x_2, x_1], x_3]$ e $[[x_3, x_1], x_2]$. Esses dois últimos polinômios constituem a base de Hall procurada.

Exemplo 1.44. *Seja $X = \{x_1, x_2, x_3, x_4\}$. Construiremos uma base de Hall para o subespaço dos polinômios de Lie multilineares em $\mathbb{K}\langle X\rangle$ de grau até 4.*

Ordenamos X pondo $x_1 < x_2 < x_3 < x_4$. Portanto, os monômios de grau 2 (já com uma ordenação) serão: $[x_2, x_1] < [x_3, x_1] < [x_3, x_2] < [x_4, x_1] < [x_4, x_2] < [x_4, x_3]$. Os monômios de grau 3 serão:

$$\begin{aligned}[][[x_2, x_1], x_3] < [[x_2, x_1], x_4] < [[x_3, x_1], x_2] < [[x_3, x_1], x_4] < [[x_3, x_2], x_4] \\ < [[x_4, x_1], x_2] < [[x_4, x_1], x_3] < [[x_4, x_2], x_3].\end{aligned}$$

E por fim obtemos os polinômios de grau 4:

$$\begin{aligned}[][[[x_2, x_1], x_3], x_4] < [[[x_3, x_1], x_2], x_4] < [[[x_4, x_1], x_2], x_3] < [[x_4, x_3], [x_2, x_1]] \\ < [[x_4, x_2], [x_3, x_1]] < [[x_4, x_1], [x_3, x_2]].\end{aligned}$$

Todos os polinômios acima formam a base de Hall desejada. Observamos ainda que, se quiséssemos apenas uma base (não necessariamente ordenada), então não precisaríamos ordenar os monômios de grau 3 e 4.

Abaixo, usaremos as notações: $[x_1, x_2, x_3] = [[x_1, x_2], x_3]$ e indutivamente $[x_1, \ldots, x_n] = [[x_1, \ldots, x_{n-1}], x_n]$.

Definição 1.45. *Um polinômio $f \in \mathbb{K}\langle X\rangle$ é dito próprio se é uma combinação linear de produtos de comutadores, ou seja,*

$$f(x_1,\ldots,x_m)=\sum \alpha_{i,\ldots,j}[x_{i_1},\ldots,x_{i_p}]\ldots[x_{j_1},\ldots,x_{j_q}],\alpha_{i,\ldots,j}\in\mathbb{K}.$$

Assumiremos que 1 é um produto de uma quantidade nula de comutadores. Denotamos por $\mathfrak{B}$ o conjunto de todos os polinômios próprios em $\mathbb{K}\langle X\rangle$.

A próxima proposição descreve uma base para $\mathfrak{B}$, cuja demonstração é baseada no Teorema de Poincaré-Birkhoff-Witt e pode ser encontrada em [9], p. 11.

Proposição 1.46. *Escolhida uma base ordenada para a álgebra de Lie livre $L(X)$, digamos,*

$$x_1,x_2,\ldots,[x_{i_1},x_{i_2}],[x_{j_1},x_{j_2}],\ldots,[x_{k_1},x_{k_2},x_{k_3}],\ldots,$$

em que as variáveis precedem os comutadores. Então, o espaço vetorial $\mathbb{K}\langle X\rangle$ possui base

$$x_1^{a_1}\cdots x_m^{a_m}[x_{i_1},x_{i_2}]^b\cdots[x_{l_1},\cdots,x_{l_p}]^c,$$

onde $a_1,\ldots,a_m,b,\ldots,c\geqslant 0$ e $[x_{i_1},x_{i_2}]<\cdots<[x_{l_1},\ldots,x_{l_p}]$ na ordenação da base de $L(X)$. Os elementos da base de $\mathbb{K}\langle X\rangle$ com $a_1=\cdots=a_m=0$ formam uma base para o espaço vetorial $\mathfrak{B}$ dos polinômios próprios.

Considerando uma base para o espaço dos polinômios próprios, obtemos o seguinte corolário.

Corolário 1.47. *Seja $car(\mathbb{K})=0$. Um polinômio $f(x_1,\ldots,x_m)\in\mathbb{K}\langle X\rangle$ é próprio se, e somente, se todas suas derivadas parciais são iguais à zero.*

Capítulo 2

IMAGENS DE POLINÔMIOS E IDEAIS DE LIE

Neste capítulo iremos estudar a estrutura de Lie nos subespaços gerados pelas imagens de polinômios não comutativos, bem como apresentar uma classificação dos ideais de Lie das álgebras centrais simples de dimensão finita. Além de alguns outros tópicos relacionados, iremos também introduzir a conjectura que motivou este trabalho.

A principal referência deste capítulo é o artigo [6].

Sejam $\mathbb{K}$ um corpo, m um inteiro positivo e $f(x_1, \ldots, x_m) \in \mathbb{K}\langle X \rangle$. Consideremos também uma $\mathbb{K}$-álgebra associativa $\mathcal{A}$ e denotemos por $\mathcal{A}^m$ o produto cartesiano $\mathcal{A} \times \cdots \times \mathcal{A}$ m vezes. Portanto, f induz a seguinte função polinomial, que, por abuso de notação, também será denotada por f,

$$\begin{array}{cccc} f: & \mathcal{A}^m & \to & \mathcal{A} \\ & (a_1, \ldots, a_m) & \mapsto & f(a_1, \ldots, a_m) \end{array}$$

O conjunto imagem dessa função será denotado por $f(\mathcal{A})$, e também será dito ser o conjunto imagem do polinômio f sobre a álgebra $\mathcal{A}$.

Se $L_1, \ldots, L_m$ são subconjuntos de $\mathcal{A}$, então $f(L_1, \ldots, L_m)$ denotará o conjunto dos valores $f(a_1, \ldots, a_m)$, em que $a_1 \in L_1, \ldots, a_m \in L_m$.

Se U e V são subespaços de $\mathcal{A}$, então $[U, V]$ denota o subespaço gerado pelos comutadores $[u, v] = uv - vu$, em que $u \in U$ e $v \in V$.

Exemplo 2.1. *O espaço das matrizes de traço zero $sl_n(\mathbb{K})$ é igual a $[M_n(\mathbb{K}), M_n(\mathbb{K})]$. De fato, notemos que o conjunto $\{e_{i,j} | i, j = 1, \ldots, n, i \neq j\} \cup \{e_{i,i} - e_{i+1,i+1} | i = 1, \ldots, n-1\}$ é uma base de $sl_n(\mathbb{K})$. Desde que $e_{i,j} = [e_{i,j}, e_{j,j}]$, para $i \neq j$, e $e_{i,i} - e_{i+1,i+1} = [e_{i,i+1}, e_{i+1,i}]$, segue que $sl_n(\mathbb{K}) \subset [M_n(\mathbb{K}), M_n(\mathbb{K})]$. A inclusão contrária segue facilmente. Portanto, a igualdade nos garante que toda matriz de traço zero pode ser escrita como combinação linear de comutadores. No próximo capítulo iremos provar algo mais forte, que toda matriz de traço zero pode ser escrita como um único comutador entre duas matrizes de $M_n(\mathbb{K})$.*

A seguir apresentamos a definição de ideal de Lie de uma álgebra associativa.

Definição 2.2. *Um subespaço L de uma álgebra $\mathcal{A}$ é um ideal de Lie de $\mathcal{A}$ se $[L, \mathcal{A}] \subset L$.*

Exemplo 2.3. *Se $\mathcal{A} = M_n(\mathbb{K})$, então $\{0\}, \mathbb{K} \cdot I_n$, e $M_n(\mathbb{K})$ são exemplos triviais de ideais de Lie de $\mathcal{A}$. O conjunto das matrizes de traço zero $sl_n(\mathbb{K})$ também é um exemplo de ideal de Lie de $\mathcal{A}$, pois $tr(AB - BA) = tr(AB) - tr(BA) = 0$, para quaisquer matrizes $A, B \in \mathcal{A}$.*

2.1 O subespaço gerado pela imagem de um polinômio não comutativo

Iniciamos esta seção com a seguinte observação: se $x_1, \ldots, x_m, x_{m+1} \in X$, então

$$\begin{aligned}[x_1 \cdots x_m, x_{m+1}] &= [x_1, x_{m+1}]x_2 \cdots x_m + x_1[x_2, x_{m+1}]x_3 \cdots x_m \\ &\quad + \cdots + x_1 \cdots x_{m-1}[x_m, x_{m+1}],\end{aligned}$$

e portanto para qualquer polinômio multilinear $h(x_1, \ldots, x_m) \in \mathbb{K}\langle X\rangle$ temos

$$\begin{aligned}[h(x_1, \ldots, x_m), x_{m+1}] &= h([x_1, x_{m+1}], x_2, \ldots, x_m) + h(x_1, [x_2, x_{m+1}], x_3, \ldots, x_m) \\ &\quad + \cdots + h(x_1, \ldots, x_{m-1}, [x_m, x_{m+1}]).\end{aligned}$$

Se V é um $\mathbb{K}$-espaço vetorial e W é um subconjunto de V, então span(W) denotará o subespaço de V gerado por W.

No próximo teorema provaremos que o subespaço gerado pela imagem de um polinômio é um ideal de Lie.

Teorema 2.4. *Sejam $\mathbb{K}$ um corpo infinito, $\mathcal{A}$ uma $\mathbb{K}$-álgebra, $L_1, \ldots, L_m$ ideais de Lie de $\mathcal{A}$ e $f(x_1, \ldots, x_m) \in \mathbb{K}\langle X\rangle$. Então, span$(f(L_1, \ldots, L_m))$ é um ideal de Lie de $\mathcal{A}$.*

Demonstração. Escreva $f = f_0 + f_1 + \cdots + f_k$, em que f_i é a componente homogênea de grau i de f na variável x_1. Logo,

$$f(\lambda a_1, a_2, \ldots, a_m) = \sum_{i=0}^{k} \lambda^i f_i(a_1, a_2, \ldots, a_m),$$

quaisquer que sejam $\lambda \in \mathbb{K}$ e $a_j \in L_j, j \in \{1, \ldots, m\}$.

Pelo Lema 1.22, cada $f_i(a_1, \ldots, a_m) \in$ span$(f(L_1, \ldots, L_m))$. Portanto, podemos supor que f é multi-homogêneo de multigrau $(k_1, \ldots, k_m)$. Segue que existe $h \in \mathbb{K}\langle X\rangle$ multilinear tal que

$$f = h(x_1, \ldots, x_1, \ldots, x_m, \ldots, x_m)$$

em que cada x_i ocorre k_i vezes.

Observamos também que $f(a_1 + \lambda a_1', a_2, \ldots, a_m)$ será uma soma da forma $\sum_{i=0}^{k_1} \lambda^i c_i$, em que $a_1, a_1' \in L_1, a_2 \in L_2, \ldots, a_m \in L_m$. Em particular,

$$\begin{aligned} c_1 &= h(a_1', a_1, \ldots, a_1, \ldots, a_m, \ldots, a_m) \\ &\quad + \cdots + h(a_1, \ldots, a_1, a_1', \ldots, a_m, \ldots, a_m). \end{aligned}$$

Novamente pelo Lema 1.22 temos que $c_1 \in \operatorname{span}(f(L_1, \ldots, L_m))$.

Para qualquer $b \in \mathcal{A}$, temos

$$\begin{aligned} [f(a_1, \ldots, a_m), b] &= h([a_1, b], a_1, \ldots, a_1, \ldots, a_m, \ldots, a_m) \\ &\quad + \cdots + h(a_1, \ldots a_1, [a_1, b], \ldots, a_m, \ldots, a_m) \\ &\quad + \cdots + h(a_1, \ldots, a_1, \ldots, [a_m, b], a_m, \ldots, a_m) \\ &\quad + \cdots + h(a_1, \ldots, a_1, \ldots, a_m, \ldots, a_m, [a_m, b]). \end{aligned} \tag{2.1}$$

Portanto, a soma das k_1 primeiras parcelas de (2.1) pertence à $\operatorname{span}(f(L_1, \ldots, L_m))$. O c_1 produzido por $f(a_1, a_2 + \lambda a_2', a_3, \ldots, a_m)$ nos garante que a soma das próximas k_2 parcelas de (2.1) também pertence à $f(L_1, \ldots, L_m)$. Continuando com esse processo, obtemos $[f(a_1, \ldots, a_m), b] \in \operatorname{span}(f(L_1, \ldots, L_m))$. □

2.2 Ideais de Lie para algumas classes de álgebras

Em [15] Herstein provou o seguinte resultado.

Proposição 2.5. *Os ideais de Lie de $M_n(\mathbb{K})$, em que $car(\mathbb{K}) \neq 2$ ou $n \neq 2$, são $\{0\}$, $\mathbb{K}$ (denotando o conjunto das matrizes escalares), $sl_n(\mathbb{K})$ ou $M_n(\mathbb{K})$.*

O próximo resultado será uma versão da proposição anterior para uma classe de álgebras que contém as álgebras de matrizes.

Teorema 2.6. *Seja $\mathcal{A}$ um $\mathbb{K}$-álgebra central simples de dimensão finita tal que $dim_{\mathbb{K}}\mathcal{A} \neq 4$ ou $car(\mathbb{K}) \neq 2$. Se L é um ideal de Lie de $\mathcal{A}$, então L é igual a $\{0\}, \mathbb{K}$ (denotando o centro de $\mathcal{A}$) , $[\mathcal{A}, \mathcal{A}]$ ou $\mathcal{A}$.*

Demonstração. Seja $\overline{\mathbb{K}}$ o fecho algébrico de $\mathbb{K}$ e considere a extensão de escalares $\overline{\mathcal{A}} = \mathcal{A} \otimes \overline{\mathbb{K}}$. Pelo Teorema 1.31, obtemos

$$\overline{\mathcal{A}} \simeq M_n(\overline{\mathbb{K}}), \text{ em que } n = \sqrt{dim_{\mathbb{K}}\mathcal{A}}.$$

Sendo $\overline{L} = L \otimes \overline{\mathbb{K}}$ um ideal de Lie de $\overline{\mathcal{A}}$, segue da Proposição 2.5 que $\overline{L}$ é igual à $\{0\}, \overline{\mathbb{K}}, [\overline{\mathcal{A}}, \overline{\mathcal{A}}]$ ou $\overline{\mathcal{A}}$. Por outro lado,

$$\{0\} = \{0\} \otimes \overline{\mathbb{K}}, \quad \overline{\mathbb{K}} = \mathbb{K} \otimes \overline{\mathbb{K}}, \quad [\overline{\mathcal{A}}, \overline{\mathcal{A}}] = [\mathcal{A}, \mathcal{A}] \otimes \overline{\mathbb{K}} \text{ e } \overline{\mathcal{A}} = \mathcal{A} \otimes \overline{\mathbb{K}},$$

e pela Proposição 1.21 temos que L é igual a $\{0\}, \mathbb{K}, [\mathcal{A}, \mathcal{A}]$ ou $\mathcal{A}$. □

Desde que podemos considerar $\mathcal{A}$ como um subespaço da sua álgebra de quocientes centrais $Q_Z(\mathcal{A})$, então dado um subespaço V de $\mathcal{A}$, denotaremos por $\hat{V}$ o subespaço gerado por V sobre $\hat{Z}$, o centro de $Q_Z(\mathcal{A})$. Em particular, $\hat{\mathcal{A}}$ será igual à $Q_Z(\mathcal{A})$.

Corolário 2.7. *Seja $\mathcal{A}$ uma PI-álgebra prima sobre $\mathbb{K}$ tal que $dim_{\hat{Z}}(\hat{\mathcal{A}}) \neq 4$ ou $car(\mathbb{K}) \neq 2$. Se L é um ideal de Lie de $\mathcal{A}$, então $\hat{L}$ é igual a $\{0\}, \hat{Z}, [\hat{\mathcal{A}}, \hat{\mathcal{A}}]$ ou $\hat{\mathcal{A}}$.*

Demonstração. Notemos que $\hat{L}$ é um ideal de Lie de $\hat{\mathcal{A}}$ e que pelo Teorema 1.40 (Posner), $\hat{\mathcal{A}}$ é um álgebra central simples de dimensão finita. Agora basta aplicar a Proposição 2.6. □

2.3 Classificação de polinômios via suas imagens

Comecemos esta seção com a seguinte definição.

Definição 2.8. *Seja $\mathbb{K}$ um corpo. Dois polinômios $f, g \in \mathbb{K}\langle X\rangle$ são ditos ciclicamente equivalentes se $f - g$ pode ser escrito como soma de comutadores em $\mathbb{K}\langle X\rangle$. Denotaremos o fato de f e g serem ciclicamente equivalentes por $f \overset{cyc}{\sim} g$. Também diremos que f é ciclicamente equivalente à uma identidade de uma álgebra $\mathcal{A}$, se $f \overset{cyc}{\sim} p$, em que $p \in T(\mathcal{A})$.*

Em seguida veremos um lema simples e o teorema principal desta seção.

Lema 2.9. *Seja $f = f(x_1, \dots, x_m) \in \mathbb{K}\langle X\rangle$. Se f é linear em x_m, então existe $g = g(x_1 \dots, x_{m-1}) \in \mathbb{K}\langle X\rangle$ tal que $f \overset{cyc}{\sim} gx_m$.*

Demonstração. Basta provar para os monômios de f. O caso mais interessante é quando $f = px_mp'$ em que p e p' são monômios em $x_1, \dots, x_{m-1}$. Portanto, notamos que

$$px_mp' - p'px_m = [px_m, p'].$$

□

Observamos que dado um polinômio $f \in \mathbb{K}\langle X\rangle$ e uma $\mathbb{K}$-álgebra $\mathcal{A}$, temos duas possibilidades: f é ciclicamente equivalente à uma identidade polinomial de $\mathcal{A}$ ou não. Para o primeiro caso, ainda podemos classificar f entre ser ou não uma identidade polinomial de $\mathcal{A}$. Já para o segundo caso, podemos classificar f em ser ou não um polinômio central de $\mathcal{A}$. Temos assim o seguinte teorema.

Teorema 2.10. *Sejam $\mathcal{A}$ uma $\mathbb{K}$-álgebra central simples de dimensão finita, $f \in \mathbb{K}\langle X\rangle$ e $L = span(f(\mathcal{A}))$. Se $car(\mathbb{K}) = 0$, então exatamente uma das seguintes possibilidades ocorre:*

(i) f é uma identidade polinomial de $\mathcal{A}$; logo $L = \{0\}$;

(ii) *f é um polinômio central de $\mathcal{A}$; logo $L = \mathbb{K}$;*

(iii) *f não é uma identidade polinomial de $\mathcal{A}$, mas é ciclicamente equivalente à uma identidade de $\mathcal{A}$; logo $L = [\mathcal{A}, \mathcal{A}]$;*

(iv) *f não é um polinômio central de $\mathcal{A}$, nem ciclicamente equivalente à uma identidade de $\mathcal{A}$; logo $L = \mathcal{A}$.*

Demonstração. Pelo Teorema 2.6 sabemos que L deve ser igual à $\{0\}, \mathbb{K}, [\mathcal{A}, \mathcal{A}]$ ou $\mathcal{A}$. Desde que $car(\mathbb{K}) = 0$ então $\overline{\mathbb{K}} \cap sl_n(\overline{\mathbb{K}}) = \{0\}$. Pelo isomorfismo $\mathcal{A} \otimes \overline{\mathbb{K}} \simeq M_n(\overline{\mathbb{K}})$, obtemos $\mathbb{K} \cap [\mathcal{A}, \mathcal{A}] = \{0\}$.

Comecemos supondo que f é ciclicamente equivalente à uma identidade de $\mathcal{A}$. Então, $f(\mathcal{A}) \subset [\mathcal{A}, \mathcal{A}]$ e, portanto, $L \subset [\mathcal{A}, \mathcal{A}]$. Sendo $\mathbb{K} \cap [\mathcal{A}, \mathcal{A}] = \{0\}$, então temos apenas duas possibilidades: $L = \{0\}$ ou $L = [\mathcal{A}, \mathcal{A}]$. Se f é uma identidade polinomial de $\mathcal{A}$, então segue o item (i), caso contrário temos o item (iii).

Supomos agora que f não é ciclicamente equivalente à uma identidade polinomial de $\mathcal{A}$. Se f é um polinômio central de $\mathcal{A}$ então o item (ii) está concluído. Portanto, assumimos que f é não central e iremos provar que $L = \mathcal{A}$. Obviamente, já temos $L \neq \{0\}$ e $L \neq \mathbb{K}$.

Suponha, por absurdo, que $L = [\mathcal{A}, \mathcal{A}]$. Segue que $f(\mathcal{A}) \subset [\mathcal{A}, \mathcal{A}]$ e, portanto, o mesmo ocorre com cada componente multi-homogênea de f. Certamente, ao menos uma dessas componentes não é ciclicamente equivalente à uma identidade de $\mathcal{A}$. Em suma, existe $h(x_1, \ldots, x_m) \in \mathbb{K}\langle X\rangle$ multi-homogêneo que não é ciclicamente equivalente à uma identidade de $\mathcal{A}$ e tal que $h(\mathcal{A}) \subset [\mathcal{A}, \mathcal{A}]$. Provaremos, porém, que tal polinômio não pode existir. Em outras palavras, provaremos que todo polinômio multi-homogêneo cuja imagem sobre $\mathcal{A}$ está contida em $[\mathcal{A}, \mathcal{A}]$ deve ser ciclicamente equivalente à uma identidade de $\mathcal{A}$.

Procedemos por indução sobre o grau k de h na variável x_m: se $k = 1$, pelo Lema 2.9, existe $g = g(x_1, \ldots, x_{m-1}) \in \mathbb{K}\langle X\rangle$ tal que $h \overset{cyc}{\sim} gx_m$. Consequentemente, $(gx_m)(\mathcal{A}) \subset [\mathcal{A}, \mathcal{A}]$. Sejam $a_1, \ldots, a_{m-1} \in \mathcal{A}$ quaisquer e escreva $\omega = g(a_1, \ldots, a_{m-1})$. Logo, $\omega a \in [\mathcal{A}, \mathcal{A}]$, para todo $a \in \mathcal{A}$. Se $\omega \neq 0$, então pela simplicidade de $\mathcal{A}$, o ideal bilateral gerado por ω é igual à $\mathcal{A}$. Logo, existem $u_i, v_i \in \mathcal{A}$ tais que $1 = \sum_i u_i \omega v_i$. Então,

$$1 = \sum_i [u_i, \omega v_i] + \omega \sum_i v_i u_i \in [\mathcal{A}, \mathcal{A}],$$

contradizendo $\mathbb{K} \cap [\mathcal{A}, \mathcal{A}] = \{0\}$. Portanto, $\omega = 0$, ou seja, g é uma identidade polinomial de $\mathcal{A}$, provando que h é ciclicamente equivalente à uma identidade de $\mathcal{A}$. Agora supomos $k > 1$ e consideramos

$$\begin{aligned} h'(x_1, \ldots, x_m, x_{m+1}) &= h(x_1, \ldots, x_{m-1}, x_m + x_{m+1}) - h(x_1, \ldots, x_{m-1}, x_m) \\ &\quad - h(x_1, \ldots, x_{m-1}, x_{m+1}). \end{aligned}$$

Logo, $h'(\mathcal{A}) \subset [\mathcal{A}, \mathcal{A}]$, e assim o mesmo ocorre com suas componentes multi-homogêneas. Como o grau de cada uma dessas componentes na variável x_m é menor que k, então cada uma delas é ciclicamente equivalente à uma identidade de $\mathcal{A}$, pela hipótese de indução. Portanto, h' também é ciclicamente equivalente à uma identidade de $\mathcal{A}$, o que encerra a prova já que

$$h'(x_1, \ldots, x_{m-1}, x_m, x_m) = (2^k - 2)h(x_1, \ldots, x_m).$$

□

Corolário 2.11. *Sejam $n \geqslant 2$, $car(\mathbb{K}) = 0$ e $f = f(x_1, \ldots, x_m) \in \mathbb{K}\langle X\rangle$. Então, $tr(f(A_1, \ldots, A_m)) = 0$, para todo $A_i \in M_n(\mathbb{K})$ se, e somente, se f é ciclicamente equivalente à uma identidade de $M_n(\mathbb{K})$.*

Demonstração. Supondo $tr(f(A_1, \ldots, A_m)) = 0$ para todo $A_i \in M_n(\mathbb{K})$, segue pelo Exemplo 2.1 que $f(M_n(\mathbb{K})) \subset [M_n(\mathbb{K}), M_n(\mathbb{K})]$. Pela demonstração do Teorema 2.10, f é ciclicamente equivalente à uma identidade de $M_n(\mathbb{K})$.

Reciprocamente, se f não é ciclicamente equivalente à uma identidade de $M_n(\mathbb{K})$, então, pelo Teorema 2.10, $\mathrm{span}(f(M_n(\mathbb{K}))) \in \{\mathbb{K}, M_n(\mathbb{K})\}$, contradizendo a hipótese. □

Corolário 2.12. *Suponha $car(\mathbb{K}) = 0$ e $f(x_1, \ldots, x_m) \in \mathbb{K}\langle X\rangle$. Então, $tr(f(A_1, \ldots, A_m)) = 0$ para todo $A_i \in M_n(\mathbb{K})$ e todo $n \geqslant 2$ se, e somente se, $f \overset{cyc}{\sim} 0$.*

Demonstração. Suponha $tr(f(A_1, \ldots, A_m)) = 0$ para todo $A_i \in M_n(\mathbb{K})$ e para todo $n \geqslant 2$. Então, pelo Corolário 2.11, temos que para cada $n \geqslant 2$, $f \overset{cyc}{\sim} p_n$, em que $p_n \in T(M_n(\mathbb{K}))$. Fixemos $n \geqslant 2$ tal que $deg(f) < 2n$. Suponha que $p_n \neq 0$. Pela Proposição 1.36, temos que todas as componentes multi-homogêneas de p_n possuem grau maior ou igual à $2n$. Sendo f igual à soma de p_n com uma soma de comutadores e $deg(f) < 2n$, segue que cada componente multi-homogênea de p_n deve ser também uma soma de comutadores. Isso prova que $p_n \overset{cyc}{\sim} 0$ e, portanto, $f \overset{cyc}{\sim} 0$. A recíproca é óbvia. □

Dado um ideal de Lie L de $M_n(\mathbb{K})$, em que $car(\mathbb{K}) = 0$, perguntamos quais polinômios em $\mathbb{K}\langle X\rangle$ possuem imagens sobre $M_n(\mathbb{K})$ contidas em L. Para os casos em que $L = \{0\}$ e $L = \mathbb{K}$, temos as identidades polinomiais e polinômios centrais, respectivamente. Quando consideramos $L = sl_n(\mathbb{K})$, o Corolário 2.11 acima responde a questão levantada.

2.4 Matrizes genéricas de traço zero

Começamos esta seção esclarecendo algumas notações que serão usadas.

Denotaremos o conjunto das sequências em $M_n(\mathbb{K})$ por $M_n(\mathbb{K})^{\mathbb{N}}$. Portanto, escrever $A \in M_n(\mathbb{K})^{\mathbb{N}}$ significa que A é da forma $(A_1, A_2, \ldots)$, $A_i \in M_n(\mathbb{K}), i \in \mathbb{N}$.

Denotando as matrizes genéricas elementares por $y_i, i \in \mathbb{N}$, então dada $A = (A_i)_{i\in\mathbb{N}} \in M_n(\mathbb{K})^{\mathbb{N}}$, em que $\mathbb{K}$ é um corpo infinito, a função $y_i \mapsto A_i$ se estende unicamente para um homomorfismo $\varphi_A : GM_n(\mathbb{K}) \to M_n(\mathbb{K})$ pela Proposição 1.39. Portanto, para cada $f \in GM_n(\mathbb{K})$, denotaremos $\varphi_A(f)$ por $f(A)$.

Teorema 2.13. *Sejam $\mathbb{K}$ um corpo de característica zero e $f \in GM_n(\mathbb{K})$. Então as seguintes afirmações são equivalentes.*

(i) f é uma soma de comutadores em $GM_n(\mathbb{K})$;

(ii) $tr(f) = 0$;

(iii) $tr(f(A)) = 0$, para todo $A \in M_n(\mathbb{K})^{\mathbb{N}}$;

(iv) $f(A)$ é uma soma de comutadores em $M_n(\mathbb{K})$, para todo $A \in M_n(\mathbb{K})^{\mathbb{N}}$.

Demonstração. As implicações $(ii) \Rightarrow (iii)$, $(i) \Rightarrow (iii)$ e $(iv) \Rightarrow (iii)$ são óbvias. As implicações $(iii) \Rightarrow (ii)$ e $(iii) \Rightarrow (iv)$ seguem pelo Lema 1.28 e pelo Exemplo 2.1, respectivamente. Portanto, basta provarmos que $(iii) \Rightarrow (i)$.

Existe um único homomorfismo $\varphi : \mathbb{K}\langle X\rangle \to GM_m(\mathbb{K})$ que estende a função $x_i \mapsto y_i$. Notemos que tal homomorfismo é sobrejetor. Portanto, para $f \in GM_n(\mathbb{K})$, existe $F(x_1, \ldots, x_m)$ tal que $f = \varphi(F(x_1, \ldots, x_m))$, ou seja, $f(A) = F(Y_1(A), \ldots, Y_m(A))$, qualquer que seja $A \in M_n(\mathbb{K})^{\mathbb{N}}$. Quando A percorre todo $M_n(\mathbb{K})^{\mathbb{N}}$, então $(Y_1(A), \ldots, Y_m(A))$ percorre todas as m-uplas com entradas em $M_n(\mathbb{K})$. Logo, $tr(F(A_1, \ldots, A_m)) = 0$ para quaisquer $A_i \in M_n(\mathbb{K})$. Pelo Corolário 2.11, F é ciclicamente equivalente à uma identidade polinomial de $M_n(\mathbb{K})$. Segue que

$$f = F(y_1, \ldots, y_m) = g(y_1, \ldots, y_m) + \sum[\cdot, \cdot]$$

para algum $g \in T(M_n(\mathbb{K}))$, e tal que $\sum[\cdot, \cdot]$ representa uma soma de comutadores em $\mathbb{K}\langle X\rangle$. Pela Proposição 1.39, segue que $g(y_1, \ldots, y_m) = 0$, e, portanto, f é uma soma de comutadores em $GM_n(\mathbb{K})$. □

2.5 A conjectura de Lvov-Kaplansky

Como exposto na introdução, um problema muito conhecido na teoria é o de determinar as imagens de polinômios multilineares sobre a álgebra das matrizes. A seguinte conjectura posta por Lvov e também atribuída à Kaplansky será o nosso principal objeto de estudo.

Conjectura 2.14 (Lvov-Kaplansky)**.** *A imagem de um polinômio multilinear em $\mathbb{K}\langle X\rangle$ sobre $M_n(\mathbb{K})$ é um espaço vetorial.*

Em [21], Malev provou que sobre corpos arbitrários, a imagem de um polinômio multilinear sobre matrizes de ordem dois será zero, as matrizes escalares ou conterá todas as matrizes de traço zero. Portanto, pela Proposição 2.5 podemos reformular a conjectura de Lvov-Kaplansky equivalentemente como:

Conjectura 2.15. *A imagem de um polinômio multilinear em $\mathbb{K}\langle X\rangle$ sobre $M_n(\mathbb{K})$ é $\{0\}, \mathbb{K}$, $sl_n(\mathbb{K})$ ou $M_n(\mathbb{K})$.*

O próximo exemplo mostra que a conjectura acima é falsa quando consideramos polinômios não multilineares.

Exemplo 2.16. *Seja $f(x) = x^n \in \mathbb{C}\langle X\rangle$. Desde que $f(e_{i,i}) = e_{i,i}$ e $f(e_{i,j}+e_{j,j}) = e_{i,j}+e_{j,j}$ para todo $i, j \in \{1, \ldots, n\}$ distintos, segue que se $f(M_n(\mathbb{K}))$ for um espaço vetorial, então devemos ter $f(M_n(\mathbb{K})) = M_n(\mathbb{K})$. Agora, seja $A \in M_n(\mathbb{K})$ uma matriz nilpotente não nula. Como 0 é o seu único autovalor, segue que a forma de Jordan de A é estritamente triangular superior, o que implica em $A^n = 0$. Por outro lado, existe $B \in M_n(\mathbb{K})$ tal que $A = f(B) = B^n$. Logo, B também é nilpotente e, pelo mesmo argumento acima, temos $B^n = 0$, ou seja, $A = 0$, o que é uma contradição. Portanto, $f(M_n(\mathbb{K}))$ não é um subespaço.*

A conjectura de Lvov-Kaplansky é trivialmente verdadeira para polinômios de grau 1. Mais ainda, a imagem de polinômios de grau 1 sobre qualquer $\mathbb{K}$-álgebra é um espaço vetorial. Por essa razão, nos próximos capítulos iremos sempre tratar de polinômios multilineares de grau maior ou igual a 2.

É notório que todo polinômio multilinear $f(x_1, \ldots, x_m) \in \mathbb{K}\langle X\rangle$ pode ser escrito como

$$f(x_1, \ldots, x_m) = \sum_{\sigma \in S_m} \alpha_\sigma x_{\sigma(1)} \cdots x_{\sigma(m)}, \alpha_\sigma \in \mathbb{K}.$$

A seguir destacamos algumas propriedades sobre imagens de polinômios multilineares sobre álgebras, que serão de fundamental importância no estudo da Conjectura 2.14 e em algumas de suas variações.

Lema 2.17. *Sejam $\mathbb{K}$ um corpo, $\mathcal{A}$ uma $\mathbb{K}$-álgebra e*

$$f(x_1, \ldots, x_m) = \sum_{\sigma \in S_m} \alpha_\sigma x_{\sigma(1)} \cdots x_{\sigma(m)} \in \mathbb{K}\langle X\rangle.$$

Sejam também $a_1, \ldots, a_m \in \mathcal{A}$. Então:

(i) Para qualquer $\alpha \in \mathbb{K}$ temos $\alpha f(a_1, \ldots, a_m) = f(\alpha a_1, \ldots, a_m)$;

(ii) Para qualquer $b \in \mathcal{A}$ invertível temos $bf(a_1, \ldots, a_m)b^{-1} = f(ba_1b^{-1}, \ldots, ba_mb^{-1})$;

(iii) Para qualquer $\tau \in S_m$*, denotando* $f_\tau := f(x_{\tau(1)}, \ldots, x_{\tau(m)})$*, temos* $f(\mathcal{A}) = f_\tau(\mathcal{A})$*;*

(iv) Se $\mathcal{A}$ *é unitária e* $\sum_{\sigma \in S_m} \alpha_\sigma \neq 0$*, então* $f(\mathcal{A}) = \mathcal{A}$*.*

Demonstração. A prova dos itens $(i), (ii)$ e (iii) são triviais. Para o item (iv), dado $A \in \mathcal{A}$ temos que $A = (\sum_{\sigma \in S_m} \alpha_\sigma)^{-1} f(A, 1, \ldots, 1) \in f(\mathcal{A})$. □

Se $f(x_1, \ldots, x_m) \in \mathbb{K}\langle X \rangle$ é multilinear, então, para cada $l \in \{1, \ldots, m\}$, a derivada parcial formal de f em relação à variável x_l (que é um polinômio multilinear) será $f(x_1, \ldots, x_{l-1}, 1, x_{l+1}, \ldots, x_m)$. Observamos que se $\mathcal{A}$ é uma $\mathbb{K}$-álgebra com unidade, então a imagem de cada derivada parcial de f sobre $\mathcal{A}$ está contida em $f(\mathcal{A})$.

Capítulo 3

SOBRE AS IMAGENS DE POLINÔMIOS MULTILINEARES DE GRAU ATÉ 4

Neste capítulo vamos classificar as imagens de polinômios multilineares de grau 2 sobre $M_n(\mathbb{K})$, em que $n \geqslant 2$. Além disso, será provado que sob algumas restrições no corpo de escalares, as imagens de polinômios multilineares de grau 3 e 4 sobre matrizes de ordem n, em que $n \geqslant 2$ e $n \geqslant 3$, respectivamente, sempre contêm as matrizes de traço zero.

3.1 Imagens de polinômios multilineares de grau 2

Nesta seção iremos classificar as imagens de polinômios multilineares de grau 2 sobre a álgebra das matrizes $M_n(\mathbb{K})$, em que $n \in \mathbb{N}$ e $\mathbb{K}$ é um corpo qualquer.

Os resultados desta seção foram extraídos do artigo [1]. Iniciamos com o seguinte lema.

Lema 3.1. *Sejam $\mathbb{K}$ um corpo qualquer e $M = (m_{i,j}) \in M_n(\mathbb{K})$ uma matriz de traço zero com $m_{i,j} = 0$ para $j \geqslant i + 2$ e*

$$\sum_{i=1}^{n-1} m_{i,i+1} = 0.$$

Então existem $A, B \in M_n(\mathbb{K})$ com A invertível tais que $M = AB - BA$.

Demonstração. Sejam $C = \sum_{j=1}^{n-1} e_{j+1,j}$, $B = (b_{i,j}) \in M_n(\mathbb{K})$ em que $b_{i,1} = 0$ para $i = 1, \ldots, n$ e $b_{i,j} = 0$ para $j \geqslant i + 3$. Então,

$$CB = \left(\sum_{j=1}^{n-1} e_{j+1,j}\right)\left(\sum_{i,j=1}^{n} b_{i,j} e_{i,j}\right) = \sum_{i=2}^{n}\sum_{j=1}^{n-1} b_{i-1,j} e_{i,j} + \sum_{i=2}^{n} b_{i-1,n} e_{i,n}$$

e

$$BC = \left(\sum_{i,j=1}^{n} b_{i,j}e_{i,j}\right)\left(\sum_{j=1}^{n-1} e_{j+1,j}\right) = \sum_{i=2}^{n}\sum_{j=1}^{n-1} b_{i,j+1}e_{i,j} + \sum_{j=1}^{n-1} b_{1,j+1}e_{1,j}.$$

Notemos que a matriz CB possui a primeira linha nula e a sua i-ésima linha é igual a $(i-1)$-ésima linha de B para $i = 2, \ldots, n$, e em relação à matriz BC, sua última coluna é nula e a sua j-ésima coluna é igual a $(j+1)$-ésima coluna de B para $j = 1, \ldots, n-1$.

Seja $H = CB - BC = (h_{i,j})$. Assim,

$$h_{1,i} = -b_{1,i+1}, i = 1, \ldots, n-1,$$

$$h_{1,n} = 0,$$

$$h_{i,n} = b_{i-1,n}, i = 2, \ldots, n,$$

$$h_{i,j} = b_{i-1,j} - b_{i,j+1}, i = 2, \ldots, n, j = 1, \ldots, n-1.$$

Note que da última equação acima já temos $h_{i,1} = -b_{i,2}, i = 2, \ldots, n$.

Para $j \geqslant i+2$, equivalentemente $j \geqslant (i-1)+3$ e $j+1 \geqslant i+3$, temos

$$h_{i,j} = b_{i-1,j} - b_{i,j+1} = 0 - 0 = 0.$$

Observamos também que

$$\sum_{i=1}^{n} h_{i,i} = \sum_{i=1}^{n-1} h_{i,i+1} = 0 \tag{3.1}$$

Fazendo $H = M$, temos

$$m_{i,1} = -b_{i,2}, i = 1 \ldots, n,$$

$$m_{1,i} = -b_{1,i+1}, i = 1, \ldots, n-1,$$

$$m_{1,n} = 0,$$

$$m_{i,n} = b_{i-1,n}, i = 2, \ldots, n,$$

$$m_{i,j} = 0, j \geqslant i+2.$$

Agora, falta determinar os $m_{i,j}$'s com $i = 2, \ldots, n$ e $j = 2, \ldots, \min\{i+1, n-1\}$ através das equações $m_{i,j} = b_{i-1,j} - b_{i,j+1}$. Observamos que para cada elemento $h_{i,j}$, em cada coluna de H, exceto a última, existe um elemento $b_{i,j}$ que não ocorre nas colunas anteriores e nem nas demais entradas da coluna que o contém. Logo, podemos determinar

B tal que a igualdade $H = M$ ocorra, já que a última coluna de H é igual a de M por (3.1). Pondo $A = C + I$, vemos que A é invertível. Daí,

$$AB - BA = (C + I)B - B(C + I) = CB + B - BC - B = CB - BC = M.$$

□

Estamos aptos a provar o resultado principal desta seção, que foi feito por Albert e Muckenhoupt em [1]. Tal resultado basicamente estuda a imagem do polinômio $f(x_1, x_2) = [x_1, x_2] \in \mathbb{K}\langle X\rangle$.

Teorema 3.2. *Sejam $\mathbb{K}$ um corpo qualquer e $M \in M_n(\mathbb{K})$ uma matriz de traço zero. Então existem $A, B \in M_n(\mathbb{K})$ tais que $M = AB - BA$.*

Demonstração. Seja N uma matriz semelhante a M. Afirmamos que M é um comutador se, e somente se, N é um comutador. De fato, se $N = P^{-1}MP = AB - BA$, então

$$M = (PAP^{-1})(PBP^{-1}) - (PBP^{-1})(PAP^{-1}) = [PAP^{-1}, PBP^{-1}],$$

e a recíproca é analogamente verdadeira.

Portanto, podemos supor que M está na forma canônica racional, ou seja,

$$M = diag(C_{\phi_1}, \dots, C_{\phi_k}),$$

em que $\phi_i = \phi_i(x)$ são os fatores invariantes não triviais de $det(xI - M)$, $\phi_i(x)$ divide $\phi_{i-1}(x)$ para $i = 2, \dots, k$ e C_ϕ é a matriz companheira

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ \alpha_m & \alpha_{m-1} & \alpha_{m-2} & \cdots & \alpha_1 \end{pmatrix} \tag{3.2}$$

do polinômio $\phi = \phi(x) = x^m - (\alpha_1 x^{m-1} + \cdots + \alpha_m)$. Desta forma, a matriz M possui apenas 0 e 1 acima da diagonal principal.

Consequentemente, existe uma conjugação que substitui todos os 1's acima da diagonal principal pela sequência $1, -1, 1, -1, \dots$. De fato, observemos que multiplicar a i-ésima linha e coluna de uma matriz por -1 é o mesmo que multiplicar à esquerda e à direita pela matriz $J_i = [e_1|\cdots| - e_i|\cdots|e_n]$, em que para cada $j = 1, \dots, n$, e_j é um vetor coluna com 1 na j-ésima linha e 0 nas demais. Sendo J_i autoinvertível, então podemos interpretar este processo como uma conjugação pela matriz J_i. Note que podemos multiplicar a 3ª linha e coluna de (3.2) por -1 e teremos no lugar do segundo 1 um -1.

Assumindo que a sequência desejada já foi produzida para os t primeiros elementos não nulos $m_{i,i+1}$ e supondo que o $(k+1)$-ésimo elemento ocorre na s-ésima linha e $(s+1)$-ésima coluna, então se necessário, multiplicamos a $(s+1)$-ésima linha e coluna por -1. Portanto, podemos supor que, a menos de conjugação, $M = (m_{i,j})$ é tal que $m_{i,j} = 0$ para $j \geqslant i+2$ com $tr(M) = 0$, $m_{i,i+1} = 1, -1$ ou 0 e os $m_{i,i+1}$'s não nulos formando uma sequência alternando em sinal.

Se houver um número par de $m_{i,i+1}$ não nulos, então temos $\sum_{i=1}^{n-1} m_{i,i+1} = 0$ e pelo Lema 3.1 temos $M = AB - BA$, como desejado.

Se o número de $m_{i,i+1}$'s não nulos for ímpar, então a matriz M tem a forma

$$M = \begin{pmatrix} 0 & u \\ v^T & M_1 \end{pmatrix},$$

em que $u, v \in M_{1\times(n-1)}(\mathbb{K})$ e $M_1 \in M_{n-1}(\mathbb{K})$. Note que M_1 possui todas as propriedades do nosso lema. Logo, $M_1 = A_1B_1 - B_1A_1$, com $A_1, B_1 \in M_{n-1}(\mathbb{K})$ e A_1 invertível.

Tomemos

$$A = \begin{pmatrix} 0 & 0 \\ 0 & A_1 \end{pmatrix} \text{ e } B = \begin{pmatrix} 0 & -uA_1^{-1} \\ A_1^{-1}v^T & B_1 \end{pmatrix}$$

e vejamos que

$$AB - BA = \begin{pmatrix} 0 & 0 \\ v^T & A_1B_1 \end{pmatrix} - \begin{pmatrix} 0 & -u \\ 0 & B_1A_1 \end{pmatrix} = \begin{pmatrix} 0 & u \\ v^T & M_1 \end{pmatrix} = M.$$

□

O teorema acima afirma que a imagem do polinômio $f(x_1, x_2) = [x_1, x_2]$ sobre $M_n(\mathbb{K})$ é igual a $sl_n(\mathbb{K})$. Ressaltamos que este teorema não vale quando consideramos matrizes com entradas em um domínio comutativo. De fato, a matriz

$$\begin{pmatrix} x & y \\ z & -x \end{pmatrix}$$

não pode ser escrita como um comutador em $\mathbb{K}[x, y, z]$. Recomendamos o artigo de Rosset e Rosset [23] para maiores detalhes.

O seguinte corolário descreve as imagens de polinômios multilineares de grau 2 sobre a álgebra das matrizes.

Corolário 3.3. *Sejam $\mathbb{K}$ um corpo qualquer e $f(x_1, x_2) \in \mathbb{K}\langle X\rangle$ multilinear. Então, $f(M_n(\mathbb{K}))$ é $\{0\}, sl_n(\mathbb{K})$ ou $M_n(\mathbb{K})$.*

Demonstração. Seja $f(x_1, x_2) = \alpha x_1x_2 + \beta x_2x_1 \in \mathbb{K}\langle X\rangle, \alpha, \beta \in \mathbb{K}$. Se $\alpha + \beta \neq 0$, então pelo Lema 2.17 (iv) temos $f(M_n(\mathbb{K})) = M_n(\mathbb{K})$. Caso contrário, podemos ter $\alpha = \beta = 0$ e, portanto, $f(M_n(\mathbb{K})) = \{0\}$, ou $\alpha = -\beta \neq 0$, o que nos fornece $f(x_1, x_2) = \alpha[x_1, x_2]$, ou seja, $f(M_n(\mathbb{K})) = sl_n(\mathbb{K})$. □

3.2 A conjectura de Mesyan

Nesta seção apresentaremos um problema relacionado à conjectura de Lvov-Kaplansky. Iniciamos com o seguinte lema técnico extraído de [2].

Proposição 3.4. *Sejam D um anel de divisão, $n \geqslant 2$ um inteiro e $A \in M_n(D)$ uma matriz não central. Então, A é semelhante a uma matriz em $M_n(D)$ com no máximo uma entrada não nula sobre a diagonal principal. Em particular, se A possui traço zero e é não central, então A é semelhante a uma matriz em $M_n(D)$ com apenas zeros na diagonal principal.*

Demonstração. A prova será feita em três casos:

1º caso: $n = 2$. Seja $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(D)$. Então,

$$\begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & -u \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a+uc & -(a+uc)u+b+ud \\ c & -cu+d \end{pmatrix} \tag{3.3}$$

Se $c \neq 0$, então escolhemos $u = -ac^{-1}$ e obtemos a forma desejada. Se $b \neq 0$, então fazemos

$$\begin{pmatrix} 1 & 0 \\ u & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -u & 1 \end{pmatrix} = \begin{pmatrix} a-bu & b \\ ua+c-(ub+d)u & ub+d \end{pmatrix}$$

e tomamos $u = b^{-1}a$.

Finalmente, supomos $b = c = 0$. Então, existe $u \in D$ tal que $-au + ud \neq 0$. De fato, se ocorresse $-au + ud = 0$, para todo $u \in D$, então, em particular, para $u = 1$ teríamos $a = d$. Daí, $au = ua$ para todo $u \in D$, o que contradiz A ser não central. Como,

$$\begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}\begin{pmatrix} 1 & -u \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & -au+ud \\ 0 & d \end{pmatrix}$$

e

$$\begin{pmatrix} 1 & 0 \\ v & 1 \end{pmatrix}\begin{pmatrix} a & -au+ud \\ 0 & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -v & 1 \end{pmatrix}$$

$$= \begin{pmatrix} a-(-au+ud)v & -au+ud \\ va-(v(-au+ud)+d)v & v(-au+ud)+d \end{pmatrix}$$

então basta escolhermos $v = (-au+ud)^{-1}a$.

2º caso: $n = 3$. Sendo A não central, uma de suas submatrizes principais é não central. Provemos que, a menos de conjugação, conseguimos escrever A tendo a sua primeira submatriz principal não central. Denotando por P_{ij} a matriz identidade com as colunas i e j trocadas entre si, então, supondo

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & k \end{pmatrix}$$

e $\begin{pmatrix} e & f \\ h & k \end{pmatrix}$ não central, temos

$$P_{23}P_{12}AP_{12}P_{23} = \begin{pmatrix} e & f & d \\ h & k & g \\ b & c & a \end{pmatrix}.$$

Caso $\begin{pmatrix} a & c \\ g & k \end{pmatrix}$ seja não central, então

$$P_{23}AP_{23} = \begin{pmatrix} a & c & b \\ g & k & h \\ d & f & e \end{pmatrix}.$$

Portanto, denotando por $A_{2\times 2}$ a primeira submatriz principal de A, que vimos que é não central, pelo 1º caso existe uma matriz $P \in M_2(D)$ invertível tal que a entrada $(1,1)$ de $PA_{2\times 2}P^{-1}$ é nula. Desde que a matriz

$$\begin{pmatrix} P & 0 \\ 0^T & 1 \end{pmatrix} A \begin{pmatrix} P^{-1} & 0 \\ 0^T & 1 \end{pmatrix}$$

possui a entrada $(1,1)$ nula, em que 0 é o vetor (coluna) nulo de $D \times D$, segue que podemos supor

$$A = \begin{pmatrix} 0 & a & b \\ c & u & v \\ d & w & z \end{pmatrix}.$$

Se tivermos $C = \begin{pmatrix} u & v \\ w & z \end{pmatrix}$ não central, então novamente pelo 1º caso, conseguimos A conjugada à uma matriz com as entradas $(1,1)$ e $(2,2)$ nulas. Suponha agora

$$A = \begin{pmatrix} 0 & a & b \\ c & \alpha & 0 \\ d & 0 & \alpha \end{pmatrix},$$

com $\alpha \in Z(D)$.

Se $c \neq 0$, então

$$\begin{pmatrix} 1 & u & v \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & a & b \\ c & \alpha & 0 \\ d & 0 & \alpha \end{pmatrix}\begin{pmatrix} 1 & -u & -v \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} uc+vd & * & * \\ c & -cu+\alpha & -cv \\ d & -du & -dv+\alpha \end{pmatrix}.$$

Tomando $u = -dc^{-1}$ e $v = 1$, então basta aplicarmos o 1º caso à submatriz principal no canto direito inferior. Caso $d \neq 0$, então basta tomar $u = 1$ e $v = -cd^{-1}$.

Para o caso $a \neq 0$ ou $b \neq 0$ consideramos a conjugação

$$\begin{pmatrix} 1 & 0 & 0 \\ u & 1 & 0 \\ v & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & a & b \\ c & \alpha & 0 \\ d & 0 & \alpha \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ -u & 1 & 0 \\ -v & 0 & 1 \end{pmatrix} = \begin{pmatrix} -au-bv & a & b \\ * & ua+\alpha & ub \\ * & va & vb+\alpha \end{pmatrix}.$$

Por último, se $a = b = c = d = 0$, temos

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix}\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 2 \\ 1 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & * & * \\ * & 0 & * \\ * & * & 2\alpha \end{pmatrix},$$

como desejado.

3º caso: $n \geqslant 3$. Por A ser não central, podemos estender analogamente os argumentos do caso anterior para garantir que exista um maior número inteiro $m \geqslant 1$ tal que A seja conjugada à

$$\begin{pmatrix} B & * \\ * & C \end{pmatrix},$$

em que é B uma matriz 0-diagonal de ordem m e C uma matriz de ordem $k = n - m$. Afirmamos que devemos ter $k = 1$, pois caso contrário poderíamos escolher a submatriz principal 3×3 formada pela última linha e coluna de B e pelas duas primeiras linhas e colunas de C. Pelo 2º caso, conseguiríamos produzir uma matriz conjugada à A começando com uma matriz 0-diagonal de ordem $m + 1$, contradizendo a maximalidade de m. □

A hipótese de que A seja uma matriz não central é de extrema importância para a validade do resultado, tendo em vista que o mesmo não vale para matrizes centrais de traço zero. De fato, considere a matriz identidade $I_2 \in M_2(\mathbb{K})$ em que $\mathbb{K}$ é um corpo de característica 2. Então I_2 possui traço zero, porém a única matriz semelhante à I_2 é ela própria, ou seja, não existem matrizes com apenas zeros na diagonal principal que sejam semelhantes à I_2.

A próxima proposição foi extraída de [22] e é uma forte motivação para conjectura de Mesyan, que será enunciada logo em seguida.

Proposição 3.5. *Sejam $\mathbb{K}$ um corpo, $n \geqslant 2$ e $m \geqslant 2$ inteiros em que $car(\mathbb{K}) \nmid n$ e $f(x_1, \ldots, x_m) \in \mathbb{K}\langle X \rangle$ um polinômio multilinear não nulo. Se $m \leqslant n+1$, então o $\mathbb{K}$-subespaço span$(f(M_n(\mathbb{K})))$ contém $sl_n(\mathbb{K})$.*

Demonstração. Escreva $f(x_1, \ldots, x_m) = \sum_{\sigma \in S_m} \alpha_\sigma x_{\sigma(1)} \ldots x_{\sigma(m)}, \alpha_\sigma \in \mathbb{K}$. Desde que f é não nulo, renomeando as variáveis se necessário, podemos assumir que $\alpha_1 \neq 0$. Além disso, pelo Corolário 3.3, temos $sl_n(\mathbb{K}) \subset f(M_n(\mathbb{K}))$ para polinômios de grau 2, para qualquer $n \geqslant 2$. Portanto, podemos assumir $m \geqslant 3$.

Sejam $i, j \in \{1, \ldots, n\}$ distintos. Como $n \geqslant m-1 \geqslant 2$, então $n-2 \geqslant m-1-2 = m-3$. Logo, podemos encontrar $m-3$ elementos distintos em $\{1, \ldots, n\} \backslash \{i, j\}$, que denotaremos por $k_1, \ldots, k_{m-3}$. Portanto,

$$\begin{aligned} f(e_{i,i}, e_{i,j}, e_{j,k_1}, e_{k_1,k_2}, \ldots, e_{k_{m-4},k_{m-3}}, e_{k_{m-3},j}) &= \alpha_1 e_{i,i} e_{i,j} e_{j,k_1} e_{k_1,k_2} \cdots e_{k_{m-4},k_{m-3}} e_{k_{m-3},j} + 0 \\ &= \alpha_1 e_{i,j}. \end{aligned}$$

Sendo $a_1 \neq 0$, concluímos que $e_{i,j} \in f(M_n(\mathbb{K}))$ para todos i, j distintos, ou seja, span$(f(M_n(\mathbb{K})))$ contém todas as matrizes cuja diagonal principal é composta por apenas zeros.

Pela Proposição 3.4, se A é uma matriz de traço zero, então por $car(\mathbb{K}) \nmid n$, temos que A é não central e daí existe $B \in \text{span}(f(M_n(\mathbb{K})))$ tal que $A = P^{-1}BP$ e usando o Lema 2.17 (ii), podemos garantir que $A \in \text{span}(f(M_n(\mathbb{K})))$. Isto prova que $sl_n(\mathbb{K}) \subset \text{span}(f(M_n(\mathbb{K})))$. □

Conjectura 3.6 (Mesyan)**.** *Sejam $\mathbb{K}$ um corpo, $n \geqslant 2$ e $m \geqslant 2$ inteiros, e $f(x_1, \ldots, x_m) \in \mathbb{K}\langle X \rangle$ um polinômio multilinear não nulo. Se $n \geqslant m-1$, então $sl_n(\mathbb{K}) \subset f(M_n(\mathbb{K}))$.*

As próximas seções responderão positivamente a conjectura de Mesyan para polinômios de grau 3 e 4.

3.3 A conjectura de Mesyan para polinômios de grau 3

Nesta seção vamos provar que sob algumas hipóteses sutis sobre o corpo $\mathbb{K}$, as imagens de polinômios multilineares de grau três em $\mathbb{K}\langle X \rangle$ sobre $M_n(\mathbb{K})$ contêm todas as matrizes de traço zero. A principal referência desta seção é o artigo [22].

Antes observamos o seguinte: a função $\partial_a : \mathcal{R} \to \mathcal{R}$ dada por $r \mapsto [a, r]$, em que $\mathcal{R}$ é uma álgebra não comutativa e $a \in \mathcal{R}$ é previamente fixado, é uma transformação linear. Notemos que $\ker \partial_a = C_{\mathcal{R}}(a) := \{r \in \mathcal{R} | ar = ra\}$ e $Im(\partial_a) = [a, \mathcal{R}]$. Portanto,

$$\dim(\mathcal{R}) = \dim(C_{\mathcal{R}}(a)) + \dim([a, \mathcal{R}]).$$

Lema 3.7. *Suponha $A = diag(\alpha_1, \ldots, \alpha_n) \in M_n(\mathbb{K})$ uma matriz diagonal com $\alpha_1, \ldots, \alpha_n$ distintos entre si. Então $[A, M_n(\mathbb{K})]$ é o conjunto das matrizes com apenas zeros na diagonal principal.*

Demonstração. Como os α_i's são distintos, então $C_{M_n(\mathbb{K})}(A)$ é igual ao conjunto das matrizes diagonais, que possui dimensão n. Pela observação acima temos que $\dim([A, M_n(\mathbb{K})]) = n^2 - n$. Seja $V = \{B = (b_{i,j}) \in M_n(\mathbb{K}) | b_{i,i} = 0, i = 1, \ldots, n\}$. Dado $C = (c_{i,j}) \in M_n(\mathbb{K})$ a entrada (i,i) de $[A, B]$ é $\alpha_i c_{i,i} - \alpha_i c_{i,i} = 0$. Logo, $[A, M_n(\mathbb{K})] \subseteq V$. Como $\dim V = n^2 - n$, temos $[A, M_n(\mathbb{K})] = V$. □

Já estamos aptos a justificar o resultado principal desta seção, que está enunciado no seguinte teorema.

Teorema 3.8. *Sejam $n \geqslant 2$ um inteiro, $\mathbb{K}$ um corpo com no mínimo n elementos tal que $car(\mathbb{K}) \nmid n$ e $f(x_1, x_2, x_3) \in \mathbb{K}\langle X \rangle$ um polinômio multilinear não nulo. Então $f(M_n(\mathbb{K}))$ contém todas as matrizes de traço zero.*

Demonstração. Seja $f(x_1, x_2, x_3) \in \mathbb{K}\langle X \rangle$ multilinear não nulo. Se $f(1, x_2, x_3), f(x_1, 1, x_3)$ ou $f(x_1, x_2, 1)$ forem não nulos, então o Corolário 3.3 nos fornece $sl_n(\mathbb{K}) \subset f(M_n(\mathbb{K}))$. Assim, assumimos

$$f(1, x_2, x_3) = f(x_1, 1, x_3) = f(x_1, x_2, 1) = 0,$$

ou seja, f é um polinômio próprio multilinear de grau 3. Isso implica em f ser um polinômio de Lie multilinear de grau 3. Pelo Exemplo 1.43, podemos escrever f como

$$f(x_1, x_2, x_3) = \alpha_1[[x_2, x_1], x_3] + \alpha_2[[x_3, x_1], x_2], \alpha_1, \alpha_2 \in \mathbb{K},$$

e sem perda de generalidade podemos supor

$$f(x_1, x_2, x_3) = [[x_2, x_1], x_3] + \alpha[[x_3, x_1], x_2], \alpha \in \mathbb{K}.$$

Agora seja $A \in M_n(\mathbb{K})$ com traço zero. Como $car(\mathbb{K}) \nmid n$, então A é não central. Queremos mostrar que $A \in f(M_n(\mathbb{K}))$. Pela Proposição 3.4, A é conjugada à uma matriz $\tilde{A} \in M_n(\mathbb{K})$ com apenas zeros na diagonal principal e pelo Lema 2.17 (ii), basta provarmos que $\tilde{A} \in f(M_n(\mathbb{K}))$. Como $\mathbb{K}$ possui pelo menos n elementos, então podemos tomar uma matriz diagonal $M \in M_n(\mathbb{K})$ que possua elementos distintos na diagonal principal. Pelo Lema 3.7, existe $B \in M_n(\mathbb{K})$ tal que $\tilde{A} = [M, B]$. Escrevendo $B = C + D$, em que C possui apenas zeros na diagonal principal e D é uma matriz diagonal, temos

$$\tilde{A} = [M, B] = [M, C + D] = [M, C] + [M, D] = [M, C].$$

Usando novamente o Lema 3.7, podemos encontrar $H \in M_n(\mathbb{K})$ tal que $C = [M, H]$, e com isso

$$f(M, H, M) = [[H, M], M] = [-C, M] = [M, C] = \tilde{A}.$$

Concluímos que $sl_n(\mathbb{K}) \subset f(M_n(\mathbb{K}))$. $\square$

Pelo teorema acima, sabemos que, sob certas condições sobre o corpo $\mathbb{K}$, a imagem de polinômios multilineares de grau 3 sobre $M_n(\mathbb{K})$ sempre contêm as matrizes de traço zero. Então uma questão natural que segue é a seguinte: com as mesmas condições sobre o corpo, quais polinômios multilineares de grau 3 possuem imagem sobre $M_n(\mathbb{K})$ exatamente igual à $sl_n(\mathbb{K})$? Para responder essa pergunta, iremos iniciar com o seguinte lema, em que usamos a notação A_3 para representar o subgrupo alternado de S_3.

Lema 3.9. *Sejam $\mathbb{K}$ um corpo, $n \geqslant 2$ um inteiro e*

$$f(x_1, x_2, x_3) = \sum_{\sigma \in S_3} \alpha_\sigma x_{\sigma(1)} x_{\sigma(2)} x_{\sigma(3)} \in \mathbb{K}\langle X \rangle.$$

Então $f(M_n(\mathbb{K})) \subset sl_n(\mathbb{K})$ se, e somente se,

$$\sum_{\sigma \in S_3} \alpha_\sigma = 0 = \sum_{\sigma \in A_3} \alpha_\sigma.$$

Demonstração. Primeiramente, suponha que $\sum_{\sigma \in S_3} \alpha_\sigma = 0 = \sum_{\sigma \in A_3} \alpha_\sigma$. Então,

$$\sum_{\sigma \in S_3 \setminus A_3} \alpha_\sigma = \sum_{\sigma \in S_3} \alpha_\sigma - \sum_{\sigma \in A_3} \alpha_\sigma = 0.$$

Portanto,

$$\begin{aligned} f(x_1, x_2, x_3) &= \sum_{\sigma \in A_3} \alpha_\sigma x_{\sigma(1)} x_{\sigma(2)} x_{\sigma(3)} + \sum_{\sigma \in S_3 \setminus A_3} \alpha_\sigma x_{\sigma(1)} x_{\sigma(2)} x_{\sigma(3)} \\ &= \sum_{\sigma \in A_3 \setminus \{(1)\}} \alpha_\sigma (x_{\sigma(1)} x_{\sigma(2)} x_{\sigma(3)} - x_1 x_2 x_3) \\ &\quad + \sum_{\sigma \in S_3 \setminus (A_3 \cup \{(12)\})} \alpha_\sigma (x_{\sigma(1)} x_{\sigma(2)} x_{\sigma(3)} - x_2 x_1 x_3) \\ &= \alpha_{(123)} (x_2 x_3 x_1 - x_1 x_2 x_3) + \alpha_{(132)} (x_3 x_1 x_2 - x_1 x_2 x_3) \\ &\quad + \alpha_{(13)} (x_3 x_2 x_1 - x_2 x_1 x_3) + \alpha_{(23)} (x_1 x_3 x_2 - x_2 x_1 x_3) \\ &= \alpha_{(123)} [x_2 x_3, x_1] + \alpha_{(132)} [x_3, x_1 x_2] + \alpha_{(13)} [x_3, x_2 x_1] + \alpha_{(23)} [x_1 x_3, x_2]. \end{aligned}$$

Logo, $f(M_n(\mathbb{K})) \subset sl_n(\mathbb{K})$.

Reciprocamente, se $\sum_{\sigma \in S_3} \alpha_\sigma \neq 0$, então o Lema 2.17 (iv) nos fornece $f(M_n(\mathbb{K})) \not\subset sl_n(\mathbb{K})$. Caso $\sum_{\sigma \in A_3} \alpha_\sigma \neq 0$, então $f(e_{11}, e_{12}, e_{21}) = \alpha_1 e_{11} + \alpha_{(123)} e_{11} + \alpha_{(132)} e_{22}$ possui traço diferente de zero, o que prova $f(M_n(\mathbb{K})) \not\subset sl_n(\mathbb{K})$. $\square$

E a resposta da pergunta levantada anteriormente é dada no seguinte corolário, cuja demonstração é uma combinação do Teorema 3.8 e do Lema 3.9.

Corolário 3.10. *Sejam $\mathbb{K}$ um corpo com no mínimo n elementos, em que $n \geqslant 2$ é um inteiro e $car(\mathbb{K}) \nmid n$, e*

$$f(x_1, x_2, x_3) = \sum_{\sigma \in S_3} \alpha_\sigma x_{\sigma(1)} x_{\sigma(2)} x_{\sigma(3)} \in \mathbb{K}\langle X \rangle$$

não nulo. Então $f(M_n(\mathbb{K})) = sl_n(\mathbb{K})$ se, e somente se,

$$\sum_{\sigma \in S_3} \alpha_\sigma = 0 = \sum_{\sigma \in A_3} \alpha_\sigma.$$

Em geral, a condição $\sum_{\sigma \in S_m} \alpha_\sigma = 0 = \sum_{\sigma \in A_m} \alpha_\sigma$ não caracteriza os polinômios multilineares não nulos tais que sua imagem coincide com o conjunto das matrizes de traço zero. Por exemplo, se $m = 2$, a condição $\sum_{\sigma \in S_2} \alpha_\sigma = 0 = \sum_{\sigma \in A_2} \alpha_\sigma$ implica em $f(x_1, x_2) = 0$ e para

$$f(x_1, x_2, x_3, x_4) = x_1 x_2 x_3 x_4 - x_4 x_3 x_2 x_1$$

temos $\sum_{\sigma \in S_4} \alpha_\sigma = 0 = \sum_{\sigma \in A_4} \alpha_\sigma$, porém $f(e_{11}, e_{12}, e_{22}, e_{21}) = e_{11} \notin sl_n(\mathbb{K})$.

3.4 A conjectura de Mesyan para polinômios de grau 4

Nesta última seção vamos apresentar o estudo feito do artigo [7] cujo resultado principal é mostrar que sobre um corpo algebricamente fechado de característica zero, a imagem de um polinômio multilinear de grau quatro avaliado na álgebra das matrizes de ordem pelo menos três contém todas as matrizes de traço zero.

Primeiramente, observamos que este resultado não é verdade para a álgebra das matrizes de ordem dois, pois os polinômios

$$f(x_1, x_2, x_3, x_4) = [x_1, x_2][x_3, x_4] + [x_3, x_4][x_1, x_2]$$

e

$$s_4(x_1, x_2, x_3, x_4) = \sum_{\sigma \in S_4} (-1)^\sigma x_{\sigma(1)} x_{\sigma(2)} x_{\sigma(3)} x_{\sigma(4)}$$

são multilineares, mas também central e identidade polinomial, respectivamente.

Iniciamos com o seguinte lema.

Lema 3.11. *Sejam $\mathbb{K}$ um corpo e $a_{i,j} \in \mathbb{K}$ elementos quaisquer com $\sum_{i=1}^{n} a_{i,i} = 0$. Seja $A = \sum_{i=1}^{n-1} e_{i,i+1} \in M_n(\mathbb{K})$. Então, existe $B \in M_n(\mathbb{K})$ tal que*

$$[A, B] = \sum_{i=1}^{n} a_{i,i} e_{i,i} + \sum_{i=1}^{n-1} a_{i,i+1} e_{i,i+1}.$$

Demonstração. Vamos escolher uma matriz B da forma $\sum_{i=1}^{n} b_{i,i}e_{i,i} + \sum_{i=1}^{n-1} b_{i+1,i}e_{i+1,i}$. Assim,

$$[A,B] = b_{2,1}e_{1,1} + \sum_{i=2}^{n-1}(b_{i+1,i} - b_{i,i-1})e_{i,i} - b_{n,n-1}e_{n,n} + \sum_{i=1}^{n-1}(b_{i+1,i+1} - b_{i,i})e_{i,i+1}.$$

Agora a prova do lema se reduz a existência de solução para os seguintes sistemas nas indeterminadas $b_{i,j}$'s:

$$\begin{cases} b_{2,1} = a_{1,1} \\ b_{3,2} - b_{2,1} = a_{2,2} \\ \quad\vdots \\ b_{n,n-1} - b_{n-1,n-2} = a_{n-1,n-1} \\ -b_{n,n-1} = a_{n,n} = -\sum_{i=1}^{n-1} a_{i,i} \end{cases} \quad \text{e} \quad \begin{cases} b_{2,2} - b_{1,1} = a_{1,2} \\ b_{3,3} - b_{2,2} = a_{2,3} \\ \quad\vdots \\ b_{n,n} - b_{n-1,n-1} = a_{n-1,n} \end{cases}.$$

É fácil ver que uma solução para os sistemas acima é

$$\begin{cases} b_{2,1} = a_{1,1} \\ b_{3,2} = a_{1,1} + a_{2,2} \\ \quad\vdots \\ b_{n,n-1} = \sum_{i=1}^{n-1} a_{i,i} \end{cases} \quad \text{e} \quad \begin{cases} b_{1,1} = 0 \\ b_{2,2} = a_{1,2} \\ b_{3,3} = a_{1,2} + a_{2,3} \\ \quad\vdots \\ b_{n,n} = \sum_{i=1}^{n-1} a_{i,i+1} \end{cases}.$$

□

Nosso objetivo a seguir é provar que, dado $\lambda \in \mathbb{K}$, a imagem do polinômio $f(x_1,x_2,x_3) = [x_1,x_2][x_1,x_3] + \lambda[x_1,x_3][x_1,x_2]$ sobre a álgebra das matrizes possui todas as matrizes de traço zero. Esse resultado é central na demonstração do resultado principal desta seção e será feito no seguinte lema.

Lema 3.12. *Sejam $\mathbb{K}$ um corpo algebricamente fechado de característica zero e $\lambda \in \mathbb{K}\backslash\{-1\}$. Então para toda matriz $D \in M_n(\mathbb{K})$ de traço zero com $n \geqslant 3$ existem $A, B, C \in M_n(\mathbb{K})$ tais que*

$$D = [A,B][A,C] + \lambda[A,C][A,B].$$

Tal Lema foi extraído de [7] e sua demonstração apresentou o seguinte erro: foi dito que sobre um corpo $\mathbb{K}$ algebricamente fechado de característica zero, para qualquer $\lambda \in \mathbb{K}\backslash\{-1\}$, o seguinte sistema nas variáveis $b_{1,1},\dots,b_{n,n}$

$$\begin{cases} (\lambda+1)b_{1,1} = d_{1,1} \\ (\lambda+1)b_{2,2} = d_{2,2} \\ \qquad \vdots \\ (\lambda+1)b_{n-1,n-1} = d_{n-1,n-1} \\ -(n-1)(\lambda+1)b_{n,n} = d_{n,n} \end{cases}$$

possuía solução satisfazendo $\sum_{i=1}^{n} b_{i,i} = 0 = \sum_{i=1}^{n} d_{i,i}$, em que $d_{i,i}$ são quaisquer elementos de $\mathbb{K}$ e n é um inteiro maior ou igual à 3. Porém isto é falso, tendo em vista que a soma do lado direito das equações do sistema acima é igual a 0 enquanto que a soma do lado esquerdo resulta em $-n(\lambda+1)b_{n,n} = 0$. Daí $b_{n,n} = 0$ o que implica em $d_{n,n} = 0$, contradizendo a generalidade de $d_{n,n}$.

No desenvolvimento desta dissertação, corrigimos esse lema. A demonstração correta é apresentada a seguir.

Demonstração. Como toda matriz com entradas num corpo algebricamente fechado é semelhante à sua forma de Jordan e, como a imagem do polinômio

$$f(x_1, x_2, x_3) = [x_1, x_2][x_1, x_3] + \lambda[x_1, x_3][x_1, x_2]$$

sobre $M_n(\mathbb{K})$ é fechada sob conjugação, podemos supor que

$$D = \sum_{i=1}^{n} d_{i,i}e_{i,i} + \sum_{i=1}^{n-1} d_{i,i+1}e_{i,i+1}. \tag{3.4}$$

Primeiramente, tome $A = \sum_{i=1}^{n-1} e_{i,i+1}$. Então, pelo Lema 3.11, dados quaisquer $a_{i,i}, b_{i,i} \in \mathbb{K}$, $i = 1, \ldots, n$ e $b_{i,i+1} \in \mathbb{K}, i = 1, \ldots, n-1$ tais que $\sum_{i=1}^{n} a_{i,i} = 0 = \sum_{i=1}^{n} b_{i,i}$, existem $B, C \in M_n(\mathbb{K})$ tais que

$$[A, B] = \sum_{i=1}^{n} a_{i,i}e_{i,i} \qquad \text{e} \qquad [A, C] = \sum_{i=1}^{n} b_{i,i}e_{i,i} + \sum_{i=1}^{n-1} b_{i,i+1}e_{i,i+1}.$$

Portanto,

$$[A, B][A, C] + \lambda[A, C][A, B] = (1+\lambda)\sum_{i=1}^{n} a_{i,i}b_{i,i}e_{i,i} + \sum_{i=1}^{n-1}(a_{i,i} + \lambda a_{i+1,i+1})b_{i,i+1}e_{i,i+1} \tag{3.5}$$

Agora a ideia é comparar as equações (3.4) e (3.5), ou seja, resolver os sistemas

$$\begin{cases} (1+\lambda)a_{1,1}b_{1,1} = d_{1,1} \\ \qquad\vdots \\ (1+\lambda)a_{n,n}b_{n,n} = d_{n,n} \end{cases} \qquad \text{e} \qquad \begin{cases} (a_{1,1}+\lambda a_{2,2})b_{1,2} = d_{1,2} \\ \qquad\vdots \\ (a_{n-1,n-1}+\lambda a_{n,n})b_{n-1,n} = d_{n-1,n} \end{cases}$$

juntamente com as condições $\sum_{i=1}^{n} a_{i,i} = 0 = \sum_{i=1}^{n} b_{i,i}$.

Inicialmente, faremos o caso particular em que todos os elementos da diagonal principal de D são nulos.

Se $\lambda \neq \frac{1}{n-1}$, vamos escolher $b_{1,1} = \cdots = b_{n,n} = 0, a_{1,1} = \cdots = a_{n-1,n-1} = 1$, $a_{n,n} = -(n-1), b_{1,2} = \frac{d_{1,2}}{1+\lambda}, b_{2,3} = \frac{d_{2,3}}{1+\lambda}, \cdots, b_{n-2,n-1} = \frac{d_{n-2,n-1}}{1+\lambda}$ e $b_{n-1,n} = \frac{d_{n-1,n}}{1+\lambda(1-n)}$.

Se $\lambda = \frac{1}{n-1} \neq 0$, então vamos escolher $b_{1,1} = \cdots = b_{n,n} = 0$, $a_{1,1} = \cdots = a_{n-2,n-2} = 1, a_{n-1,n-1} = 0, a_{n,n} = -(n-2), b_{1,2} = \frac{d_{1,2}}{1+\lambda}, \cdots$, $b_{n-3,n-2} = \frac{d_{n-3,n-2}}{1+\lambda}, b_{n-2,n-1} = d_{n-2,n-1}$ e $b_{n-1,n} = \frac{d_{n-1,n}}{\lambda(2-n)}$.

Por todos os cálculos que virão pela frente, vamos denotar $a_{i,i}$ e $b_{i,i}$ simplesmente por a_i e b_i, respectivamente. Tendo em vista os elementos da diagonal principal das matrizes em (3.4) e (3.5), queremos encontrar $a_i, b_i \in \mathbb{K}$ tais que $(1+\lambda)a_ib_i = d_{i,i}$, para cada $i \in \{1,\ldots,n\}$, além de $\sum_{i=1}^{n} a_i = 0 = \sum_{i=1}^{n} b_i$. Para nos ajudar na procura de tais elementos, vamos supor $a_i \neq 0$ para todo $i \in \{1,\ldots,n\}$. Então $b_i = (1+\lambda)^{-1}a_i^{-1}d_{i,i}$ e portanto $(1+\lambda)^{-1}(\sum_{i=1}^{n} a_i^{-1}d_{i,i}) = 0$, ou seja,

$$\frac{d_{1,1}}{a_1} + \cdots + \frac{d_{n,n}}{a_n} = 0.$$

Dividiremos a demonstração do caso geral em duas partes: a primeira será supor que a forma canônica de Jordan de D possui exatamente dois blocos de Jordan e a segunda com pelo menos três blocos de Jordan. Notemos que o caso de apenas um bloco de Jordan nos obriga termos 0 como único autovalor, já que D possui traço zero, porém este caso já foi analisado.

Suponha D com um bloco de tamanho m_1 e autovalor d_1 e outro bloco de tamanho m_2 e autovalor d_2. Existe pelo menos um bloco de tamanho 2, pois caso contrário D seria uma matriz de ordem 2. Note que podemos supor m_1 maior que ou igual a 2. Como D possui traço zero, então $d_1 = 0$ se e somente se $d_2 = 0$. Porém, o caso com apenas zeros na diagonal principal já foi resolvido, logo podemos supor $d_1 \neq 0$. Portanto, fazendo $a_3 = \cdots = a_n = 1$, obtemos

$$0 = \frac{d_1}{a_1} + \cdots + \frac{d_1}{a_{m_1}} + \frac{d_2}{a_{m_1+1}} + \cdots + \frac{d_2}{a_n} = \frac{d_1}{a_1} + \frac{d_1}{-a_1 - (n-2)} + (m_1 - 2)d_1 - m_1 d_1$$

e, portanto,

$$\frac{1}{a_1} - \frac{1}{a_1 + (n-2)} - 2 = 0, \text{ ou seja, } a_1 + (n-2) - a_1 - 2a_1^2 - 2(n-2)a_1 = 0,$$

encontrando assim a seguinte equação

$$2a_1^2 + 2(n-2)a_1 - (n-2) = 0.$$

Como $n - 2 \neq 0$, então $a_1 \neq 0$. Também temos $a_1 \neq -(n-2)$, pois caso contrário, $2(n-2)^2 + 2(n-2)(-(n-2)) - (n-2) = 0$ o que implica $2(n-2)^2 - 2(n-2)^2 - (n-2) = 0$, ou seja, $n = 2$, o que não pode ocorrer.

Notemos que as raízes da última equação são

$$a_1 = \frac{-(n-2) \pm \sqrt{n(n-2)}}{2}.$$

Sendo $a_2 = -a_1 - (n-2)$, temos

$$a_2 = \frac{-(n-2) \mp \sqrt{n(n-2)}}{2}.$$

Na determinação dos $b_{i,i+1}$'s, $b_{1,2}$ e $b_{2,3}$ dependem de a_1 e a_2 pelas equações

$$(a_1 + \lambda a_2)b_{1,2} = d_{1,2} \text{ e } (a_2 + \lambda)b_{2,3} = d_{2,3}.$$

Inicialmente, vamos escolher

$$a_1 = \frac{-(n-2) + \sqrt{n(n-2)}}{2} \text{ e } a_2 = \frac{-(n-2) - \sqrt{n(n-2)}}{2}.$$

Para $\lambda \neq -\dfrac{a_1}{a_2}$ e $\lambda \neq -a_2$, conseguimos escolher $b_{1,2}$ e $b_{2,3}$. Para $\lambda \neq -\dfrac{a_1}{a_2}$ e $\lambda = -a_2$, vamos tomar $\tilde{a_1} = a_2$ e $\tilde{a_2} = a_1$. Então na primeira equação teremos $a_2 - a_2 a_1$ que é diferente de zero pois caso contrário a_1 seria igual a 1 e na segunda equação teremos $a_1 - a_2$ que também é diferente de zero, resolvendo assim este caso. O último caso que devemos considerar é $\lambda = -\dfrac{a_1}{a_2}$. Novamente, tomando $\tilde{a_1} = a_2$ e $\tilde{a_2} = a_1$. Daí $a_2 + \lambda a_1 \neq 0$, pois caso contrário teríamos $a_1/a_2 = a_2/a_1$, i.e., $a_1^2 = a_2^2$, o que não ocorre, e $a_1 - a_1/a_2 \neq 0$, pois caso contrário $a_2 = 1$.

Como todos os λ's diferentes de -1 foram considerados na análise acima, isto encerra a prova da primeira parte.

Agora suponha que D está na forma de Jordan com $k \geqslant 3$ blocos de tamanho m_k cada. Para a matriz $[A, B] = \sum_{i=1}^{n} a_{i,i}e_{i,i}$, vamos considerar a mesma divisão em blocos que ocorre em D e num mesmo bloco iremos tomar todos os $a_{i,i}$'s iguais. Para cada $j \in \{1, \ldots, k\}$, denotaremos o elemento na diagonal principal do j-ésimo bloco de $[A, B]$ por a_j.

Suponhamos $a_j \neq 0$ para todo $j \in \{1, \ldots, k\}$. Como devemos ter $(1+\lambda)a_{i,i}b_{i,i} = d_{i,i}$, então $b_{i,i} = (1+\lambda)^{-1}a_{i,i}^{-1}d_{i,i}$. Sendo $\sum_{i=1}^{n} b_{i,i} = 0$, segue que $\sum_{i=1}^{n} a_{i,i}^{-1}d_{i,i} = 0$. Portanto,

$$\begin{aligned}
0 &= \frac{m_1 d_1}{a_1} + \cdots + \frac{m_{k-1}d_{k-1}}{a_{k-1}} + \frac{m_k d_k}{a_k} \\
&= \frac{m_1 d_1}{a_1} + \cdots + \frac{m_{k-1}d_{k-1}}{a_{k-1}} + \frac{-\sum_{j=1}^{k-1} m_j d_j}{-\sum_{j=1}^{k-1} \frac{m_j}{m_k} a_j} \\
&= \frac{m_1 d_1}{a_1} + \cdots + \frac{m_{k-1}d_{k-1}}{a_{k-1}} + \frac{m_k \sum_{j=1}^{k-1} m_j d_j}{\sum_{j=1}^{k-1} m_j a_j}.
\end{aligned}$$

Fazendo $a_1 = \cdots = a_{k-2} = 1$, obtemos

$$0 = m_1 d_1 + \cdots + m_{k-2}d_{k-2} + \frac{m_{k-1}d_{k-1}}{a_{k-1}} + \frac{m_k \sum_{j=1}^{k-1} m_j d_j}{\sum_{j=1}^{k-2} m_j + m_{k-1}a_{k-1}},$$

e, desta forma,

$$0 = a_{k-1}(\sum_{j=1}^{k-2} m_j + m_{k-1}a_{k-1})(\sum_{j=1}^{k-2} m_j d_j) + m_{k-1}d_{k-1}(\sum_{j=1}^{k-2} m_j + m_{k-1}a_{k-1}) + a_{k-1}(m_k \sum_{j=1}^{k-1} m_j d_j)$$

$$\begin{aligned}
&= m_{k-1}(\sum_{j=1}^{k-2} m_j d_j)a_{k-1}^2 + ((\sum_{j=1}^{k-2} m_j)(\sum_{j=1}^{k-2} m_j d_j) + m_{k-1}^2 d_{k-1} + m_k \sum_{j=1}^{k-1} m_j d_j)a_{k-1} \\
&\quad + m_{k-1}d_{k-1}\sum_{j=1}^{k-2} m_j.
\end{aligned}$$

Denotando $d = \sum_{j=1}^{k-2} m_j d_j$, obtemos a seguinte equação do 2º grau em a_{k-1}:

$$m_{k-1} d a_{k-1}^2 + (d \sum_{j=1}^{k-2} m_j + m_k d + m_{k-1} d_{k-1}(m_{k-1} + m_k)) a_{k-1} + m_{k-1} d_{k-1} \sum_{j=1}^{k-2} m_j = 0 \quad (3.6)$$

Como queremos que $a_{k-1} \neq 0$ e $a_k \neq 0$, e sendo $a_k = \dfrac{-\sum_{j=1}^{k-2} m_j - m_{k-1} a_{k-1}}{m_k}$, então iremos procurar por uma solução de (3.6) não nula e diferente de $-\dfrac{\sum_{j=1}^{k-2} m_j}{m_{k-1}}$. Analisemos os seguintes casos em relação aos coeficientes da equação acima:

1º caso: $m_{k-1} d \neq 0$ e $m_{k-1} d_{k-1} \sum_{j=1}^{k-2} m_j \neq 0$.

Aqui já temos duas soluções não nulas. Agora vamos provar que ao menos uma destas é diferente de $-\dfrac{\sum_{j=1}^{k-2} m_j}{m_{k-1}}$. Suponha, por absurdo, que a equação (3.6) possua duas raízes iguais a $-\sum_{j=1}^{k-2} \dfrac{m_j}{m_{k-1}}$, ou seja,

$$(a_{k-1} + \sum_{j=1}^{k-2} \frac{m_j}{m_{k-1}})^2 = 0. \quad (3.7)$$

Então $a_{k-1} = -\sum_{j=1}^{k-2} \dfrac{m_j}{m_{k-1}}$ e por outro lado, também temos

$$\begin{aligned} a_{k-1} &= -\frac{d \sum_{j=1}^{k-2} m_j + m_k d + m_{k-1} d_{k-1}(m_{k-1} + m_k)}{2 m_{k-1} d} \\ &= -\sum_{j=1}^{k-2} \frac{m_j}{2 m_{k-1}} - \frac{m_k}{2 m_{k-1}} - \frac{d_{k-1}}{2d}(m_{k-1} + m_k). \end{aligned}$$

Portanto,

$$\frac{d_{k-1}}{2d}(m_{k-1} + m_k) = -\frac{m_k}{2 m_{k-1}} + \frac{1}{2} \sum_{j=1}^{k-2} \frac{m_j}{m_{k-1}}, \text{i.e.,}$$

$$\frac{d_{k-1}}{d} = \frac{\sum_{j=1}^{k-2} m_j - m_k}{m_{k-1}(m_{k-1} + m_k)}. \quad (3.8)$$

Dividindo a equação (3.6) por $m_{k-1} d$ e comparando com a equação (3.7), obtemos

$$\begin{cases} \displaystyle\sum_{j=1}^{k-2} \frac{m_j}{m_{k-1}} + \frac{m_k}{m_{k-1}} + \frac{d_{k-1}}{d}(m_{k-1} + m_k) = 2 \sum_{j=1}^{k-2} \frac{m_j}{m_{k-1}} \\ \displaystyle\frac{d_{k-1}}{d} \sum_{j=1}^{k-2} m_j = \frac{1}{m_{k-1}^2} (\sum_{j=1}^{k-2} m_j)^2. \end{cases}$$

Como $\sum_{j=1}^{k-2} m_j \neq 0$, então $\frac{d_{k-1}}{d} = \frac{1}{m_{k-1}^2}(\sum_{j=1}^{k-2} m_j)$ e, assim, conseguimos

$$2m_{k-1}\frac{d_{k-1}}{d} = \sum_{j=1}^{k-2} \frac{m_j}{m_{k-1}} + \frac{m_k}{m_{k-1}} + \frac{d_{k-1}}{d}(m_{k-1} + m_k)$$

$$= \frac{n - m_{k-1}}{m_{k-1}} + \frac{d_{k-1}}{d}(m_{k-1} + m_k),$$

o que implica em

$$\frac{d_{k-1}}{d}(m_{k-1} - m_k) = \frac{n - m_{k-1}}{m_{k-1}},$$

e pela equação (3.8), temos

$$\frac{n - m_{k-1}}{m_{k-1}} = \frac{\sum_{j=1}^{k-2} m_j - m_k}{m_{k-1}(m_{k-1} + m_k)}(m_{k-1} - m_k), \text{ ou seja,}$$

$$n - m_{k-1} = \frac{n - m_{k-1} - 2m_k}{m_{k-1} + m_k}(m_{k-1} - m_k).$$

Portanto, devemos ter

$$(n - m_{k-1})(m_{k-1} + m_k) = (n - m_{k-1} - 2m_k)(m_{k-1} - m_k),$$

e desenvolvendo os produtos em ambos os lados da equação acima obtemos

$$nm_{k-1} + nm_k - m_{k-1}^2 - m_{k-1}m_k = nm_{k-1} - nm_k - m_{k-1}^2 + m_{k-1}m_k - 2m_{k-1}m_k + 2m_k^2,$$

que implica em

$$m_k(n - m_k) = 0, \text{ isto é, } m_k = 0 \text{ ou } n = m_k,$$

o que em ambos os casos nos fornece um absurdo.

Concluímos que existe uma raiz de (3.6) não nula e diferente de $-\sum_{j=1}^{k-2} \frac{m_j}{m_{k-1}}$.

2º caso: $m_{k-1}d \neq 0$ e $m_{k-1}d_{k-1}\sum_{j=1}^{k-2} m_j = 0$.

Como m_{k-1} e $\sum_{j=1}^{k-2} m_j$ são não nulos, então $d_{k-1} = 0$. Daí a equação (3.6) se resume à

$$m_{k-1}da_{k-1}^2 + d(\sum_{j=1}^{k-2} m_j + m_k)a_{k-1} = 0.$$

Uma solução da equação acima é $a_{k-1} = -\dfrac{\sum_{j=1}^{k-2} m_j + m_k}{m_{k-1}} \neq 0$, que também é diferente de $-\sum_{j=1}^{k-2} \dfrac{m_j}{m_{k-1}}$, pois caso contrário teríamos $m_k = 0$, o que não pode ocorrer.

3º caso: $m_{k-1}d = 0$.

Aqui devemos ter $d = 0$ e portanto a equação (3.6) torna-se

$$d_{k-1}((m_{k-1} + m_k)a_{k-1} + \sum_{j=1}^{k-2} m_j) = 0.$$

Se $d_{k-1} = 0$, então qualquer elemento de $\mathbb{K}$ é solução e portanto escolhemos a adequada.

Se $d_{k-1} \neq 0$, então $a_{k-1} = -\dfrac{\sum_{j=1}^{k-2} m_j}{m_{k-1} + m_k}$. Notemos que $a_{k-1} \neq 0$ e $a_{k-1} \neq -\sum_{j=1}^{k-2} \dfrac{m_j}{m_{k-1}}$, pois, caso contrário, $m_k = 0$.

Agora basta determinarmos os $b_{i,i+1}$'s.

Em D, na passagem do bloco que contém o elemento $d_{i,i}$ para o que contém $d_{i+1,i+1}$ devemos ter $d_{i,i+1} = 0$ e, com isso, podemos tomar $b_{i,i+1} = 0$. Para os demais elementos acima da diagonal principal teremos $(a_{l,l} + \lambda a_{l+1,l+1})b_{l,l+1} = d_{l,l+1}$ e $a_{l,l} = a_{l+1,l+1}$. Portanto, podemos tomar $b_{l,l+1} = a_{l,l}^{-1}(1+\lambda)^{-1}d_{l,l+1}$. Encerramos assim a prova deste lema. $\square$

No próximo lema iremos obter um resultado análogo ao do Lema 3.12, porém para $\lambda = -1$.

Lema 3.13. *Seja $\mathbb{K}$ um corpo algebricamente fechado de característica zero. Então para toda matriz $D \in M_n(\mathbb{K})$ de traço zero com $n \geqslant 3$, existem $A, B, C \in M_n(\mathbb{K})$ tais que $D = [[A,B],[A,C]] = [A,B][A,C] - [A,C][A,B]$.*

Demonstração. Sejam $A = \sum_{i=1}^{n-1} e_{i,i+1}$, $B = \sum_{i=1}^{n} b_{i,i}e_{i,i} + \sum_{i=1}^{n-1} b_{i+1,i}e_{i+1,i}$ com $\sum_{i=1}^{n} b_{i,i} = 0$ e $C = \sum_{i=3}^{n} (i-2)e_{i,i-2}$. Notemos que

$$[A,C] = \sum_{i=1}^{n-2} e_{i+1,i} - (n-2)e_{n,n-1}$$

e

$$[A,B] = \sum_{i=1}^{n-1} (b_{i+1,i+1} - b_{i,i})e_{i,i+1} + b_{2,1}e_{1,1} + \sum_{i=2}^{n-1} (b_{i+1,i} - b_{i,i-1})e_{i,i} - b_{n,n-1}e_{n,n}.$$

Dados $c_{i,j} \in \mathbb{K}$ com $\sum_{i=1}^{n} c_{i,i} = 0$ queremos ser capazes de escolher B tal que $[A, B] = \sum_{i=1}^{n} c_{i,i}e_{i,i} + \sum_{i=1}^{n-1} c_{i,i+1}e_{i,i+1}$. Para conseguimos B com esta propriedade devemos resolver os seguintes sistemas:

$$\begin{cases} b_{2,2} - b_{1,1} = c_{1,2} \\ b_{3,3} - b_{2,2} = c_{2,3} \\ \quad\vdots \\ b_{n,n} - b_{n-1,n-1} = c_{n-1,n} \\ b_{1,1} + \cdots + b_{n,n} = 0 \end{cases} \quad \text{e} \quad \begin{cases} b_{2,1} = c_{1,1} \\ b_{3,2} - b_{2,1} = c_{2,2} \\ \quad\vdots \\ b_{n,n-1} - b_{n-1,n-2} = c_{n-1,n-1} \\ -b_{n,n-1} = c_{n,n} \end{cases}$$

O segundo sistema possui solução

$$\begin{cases} b_{2,1} = c_{1,1} \\ b_{3,2} = c_{1,1} + c_{2,2} \\ \quad\vdots \\ b_{n,n-1} = c_{1,1} + \cdots + c_{n-1,n-1} = -c_{n,n}. \end{cases}$$

Agora observe que podemos reescrever o primeiro sistema como

$$\begin{pmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & & \\ & & & -1 & 1 \\ 1 & 1 & \cdots & 1 & 1 \end{pmatrix} \begin{pmatrix} b_{1,1} \\ b_{2,2} \\ \vdots \\ b_{n-1,n-1} \\ b_{n,n} \end{pmatrix} = \begin{pmatrix} c_{1,2} \\ c_{2,3} \\ \vdots \\ c_{n-1,n} \\ 0 \end{pmatrix}$$

Afirmamos que a matriz $n \times n$ acima é invertível. De fato, sendo quadrada basta termos o vetor nulo como o único elemento de seu núcleo.

Se

$$\begin{pmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & & \\ & & & -1 & 1 \\ 1 & 1 & \cdots & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix},$$

então

$$\begin{cases} x_2 - x_1 = 0 \\ x_3 - x_2 = 0 \\ \quad\vdots \\ x_n - x_{n-1} = 0 \\ x_1 + \cdots + x_n = 0 \end{cases},$$

que possui única solução $x_1 = \cdots = x_n = 0$.

Portanto, $[A,B] = \sum_{i=1}^{n} c_{i,i}e_{i,i} + \sum_{i=1}^{n-1} c_{i,i+1}e_{i,i+1}$ e, com isso,

$$[[A,B],[A,C]] = c_{1,2}e_{1,1} + \sum_{i=2}^{n-2}(c_{i,i+1} - c_{i-1,i})e_{i,i} + (-(n-2)c_{n-1,n} - c_{n-2,n-1})e_{n-1,n-1}$$

$$+(n-2)c_{n-1,n}e_{n,n} + \sum_{i=1}^{n-2}(c_{i+1,i+1} - c_{i,i})e_{i+1,i} - (n-2)(c_{n,n} - c_{n-1,n-1})e_{n,n-1}.$$

Agora, dada uma matriz $D \in M_n(\mathbb{K})$ de traço zero, podemos supor que $D = \sum_{i=1}^{n} d_{i,i}e_{i,i} + \sum_{i=1}^{n-1} d_{i+1,i}e_{i+1,i}$, já que $\mathbb{K}$ é algebricamente fechado. Para termos $D = [[A,B],[A,C]]$, devemos garantir a existência de soluções dos seguintes sistemas:

$$\begin{cases} c_{2,2} - c_{1,1} = d_{2,1} \\ c_{3,3} - c_{2,2} = d_{3,2} \\ \qquad \vdots \\ c_{n-1,n-1} - c_{n-2,n-2} = d_{n-1,n-2} \\ -(n-2)(c_{n,n} - c_{n-1,n-1}) = d_{n,n-1} \\ c_{1,1} + \cdots + c_{n,n} = 0 \end{cases} \quad \text{e} \quad \begin{cases} c_{1,2} = d_{1,1} \\ c_{2,3} - c_{1,2} = d_{2,2} \\ \qquad \vdots \\ c_{n-2,n-1} - c_{n-3,n-2} = d_{n-2,n-2} \\ -(n-2)c_{n-1,n} - c_{n-2,n-1} = d_{n-1,n-1} \\ (n-2)c_{n-1,n} = d_{n,n} \end{cases}.$$

O primeiro possui solução analogamente como no último sistema analisado neste lema. Já o segundo possui a seguinte solução

$$\begin{cases} c_{1,2} = d_{1,1} \\ c_{2,3} = d_{1,1} + d_{2,2} \\ \qquad \vdots \\ c_{n-2,n-1} = d_{1,1} + \cdots + d_{n-2,n-2} \\ c_{n-1,n} = \dfrac{d_{1,1} + \cdots + d_{n-1,n-1}}{-(n-2)} = \dfrac{-d_{n,n}}{-(n-2)} = \dfrac{d_{n,n}}{n-2} \end{cases}.$$

□

Lema 3.14. *Sejam $n \geqslant 3$ um inteiro, $\mathbb{K}$ um corpo de característica zero e $a_{i,j} \in \mathbb{K}$ em que $\sum_{i=1}^{n} a_{i,i} = 0$. Então existem $A,B,C \in M_n(\mathbb{K})$ tais que $[A,[[A,B],[A,C]]] = \sum_{i=1}^{n} a_{i,i}e_{i,i} + \sum_{i=1}^{n-1} a_{i,i+1}e_{i,i+1}$.*

Demonstração. No lema anterior, para uma matriz $D = \sum_{i=1}^{n} d_{i,i}e_{i,i} + \sum_{i=1}^{n-1} d_{i+1,i}e_{i+1,i}$ de traço zero, temos $D = [[A,B],[A,C]]$ e B possui a mesma forma que a matriz D. Ainda neste lema, provamos que dada uma matriz de traço zero com elementos não nulos apenas possivelmente na diagonal principal e na diagonal acima da principal pode ser escrita como $[A, B]$ com a devida escolha da matriz B. Portanto, com a devida escolha dos $d_{i,i}$'s e dos $d_{i+1,i}$'s podemos escrever

$$[A, [[A, B], [A, C]]] = [A, D] = \sum_{i=1}^{n} a_{i,i}e_{i,i} + \sum_{i=1}^{n-1} a_{i,i+1}e_{i,i+1}.$$

□

Mais um resultado parcial antes do nosso teorema principal, provaremos que as matrizes de traço zero estão na imagem do polinômio standard de grau quatro.

Proposição 3.15. *Sejam $n \geqslant 3$ inteiro e $\mathbb{K}$ um corpo algebricamente fechado de característica zero. Se*

$$\begin{aligned} f(x_1, x_2, x_3, x_4) &= [x_1, x_2][x_3, x_4] + [x_3, x_4][x_1, x_2] + [x_2, x_3][x_1, x_4] + [x_1, x_4][x_2, x_3] \\ &\quad -[x_1, x_3][x_2, x_4] - [x_2, x_4][x_1, x_3] \end{aligned}$$

então $sl_n(\mathbb{K}) \subset f(M_n(\mathbb{K}))$.

Demonstração. Sejam $A, B, C \in M_n(\mathbb{K})$. Usando a identidade $[xy, z] = x[y, z] + [x, z]y$ obtemos,

$$\begin{aligned} f(A, A^2, B, C) &= [A, A^2][B, C] + [B, C][A, A^2] + [A^2, B][A, C] + [A, C][A^2, B] \\ &\quad -[A, B][A^2, C] - [A^2, C][A, B] \\ &= [A^2, B][A, C] + [A, C][A^2, B] - [A, B][A^2, C] - [A^2, C][A, B] \\ &= (A[A, B] + [A, B]A)[A, C] + [A, C](A[A, B] + [A, B]A) \\ &\quad -[A, B](A[A, C] + [A, C]A) - (A[A, C] + [A, C]A)[A, B] \\ &= A[A, B][A, C] + [A, C][A, B]A - [A, B][A, C]A - A[A, C][A, B] \\ &= [A, [[A, B], [A, C]]]. \end{aligned}$$

Pelos Lemas 2.17 e 3.14, por f ser multilinear e considerando a forma de Jordan de uma matriz $D \in sl_n(\mathbb{K})$, temos que $sl_n(\mathbb{K}) \subset f(M_n(\mathbb{K}))$. □

Finalmente estamos prontos para demonstrar o resultado principal.

Teorema 3.16. *Sejam $n \geqslant 3$ um inteiro, $\mathbb{K}$ um corpo algebricamente fechado de característica zero e $f(x_1, x_2, x_3, x_4) \in \mathbb{K}\langle X\rangle$ um polinômio multilinear. Então, $sl_n(\mathbb{K}) \subset f(M_n(\mathbb{K}))$.*

Demonstração. Seja $f(x_1, x_2, x_3, x_4) = \sum_{\sigma \in S_4} \alpha_\sigma x_{\sigma(1)}x_{\sigma(2)}x_{\sigma(3)}x_{\sigma(4)} \in \mathbb{K}\langle X\rangle$ um polinômio multilinear de grau quatro.

Se $f(1, x_2, x_3, x_4) \neq 0, f(x_1, 1, x_3, x_4) \neq 0, f(x_1, x_2, 1, x_4) \neq 0$ ou $f(x_1, x_2, x_3, 1) \neq 0$, então pelo Teorema 3.8, temos $sl_n(\mathbb{K}) \subset f(M_n(\mathbb{K}))$. Por isso vamos supor

$$f(1, x_2, x_3, x_4) = f(x_1, 1, x_3, x_4) = f(x_1, x_2, 1, x_4) = f(x_1, x_2, x_3, 1) = 0.$$

Pelo Corolário 1.47, um tal polinômio é próprio. Assim podemos escrever

$$f(x_1, x_2, x_3, x_4) = L(x_1, x_2, x_3, x_4) + \alpha_{1234}[x_1, x_2][x_3, x_4] + \alpha_{1324}[x_1, x_3][x_2, x_4]$$

$$+\alpha_{1423}[x_1, x_4][x_2, x_3] + \alpha_{2314}[x_2, x_3][x_1, x_4] + \alpha_{2413}[x_2, x_4][x_1, x_3] + \alpha_{3412}[x_3, x_4][x_1, x_2],$$

em que $\alpha_{1234}, \alpha_{1324}, \alpha_{1423}, \alpha_{2314}, \alpha_{2413}, \alpha_{3412} \in \mathbb{K}$ e L é um polinômio de Lie em quatro variáveis.

Utilizando o Exemplo 1.44, podemos reescrever o polinômio de Lie L acima como

$$L(x_1, x_2, x_3, x_4) = \alpha_1[[[x_2, x_1], x_3], x_4] + \alpha_2[[[x_3, x_1], x_2], x_4] + \alpha_3[[[x_4, x_1], x_2], x_3]$$

$$+\alpha_4[[x_4, x_1], [x_3, x_2]] + \alpha_5[[x_4, x_2], [x_3, x_1]] + \alpha_6[[x_4, x_3], [x_2, x_1]],$$

com $\alpha_i \in \mathbb{K}, i = 1, \ldots, 6$.

As três últimas parcelas de L podem ser escritas como combinação linear de produtos de comutadores. Então podemos assumir

$$f(x_1, x_2, x_3, x_4) = \alpha_1[[[x_2, x_1], x_3], x_4] + \alpha_2[[[x_3, x_1], x_2], x_4] + \alpha_3[[[x_4, x_1], x_2], x_3]$$

$$+\alpha_{1234}[x_1, x_2][x_3, x_4] + \alpha_{1324}[x_1, x_3][x_2, x_4] + \alpha_{1423}[x_1, x_4][x_2, x_3] + \alpha_{2314}[x_2, x_3][x_1, x_4]$$

$$+\alpha_{2413}[x_2, x_4][x_1, x_3] + \alpha_{3412}[x_3, x_4][x_1, x_2].$$

Se $\alpha_1 \neq 0$, então tomemos $x_1 = x_3 = x_4 = S$, em que S é uma matriz diagonal com entradas distintas na diagonal principal. Pelo Lema 3.7, sabemos que $[M_n(\mathbb{K}), S]$ é igual ao conjunto das matrizes com apenas zeros na diagonal principal e denotaremos este último conjunto por V. Provemos que também vale $[V, S] = V$. A inclusão $[V, S] \subset V$ segue do lema referido. Se $X = (a_{i,j}) \in V$ e supondo $S = diag(b_1, \ldots, b_n)$, observemos que

$$[X, S] = \begin{pmatrix} 0 & b_2 a_{1,2} & \cdots & b_n a_{1,n} \\ b_1 a_{2,1} & 0 & \cdots & b_n a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_1 a_{n,1} & b_2 a_{n,2} & \cdots & 0 \end{pmatrix} - \begin{pmatrix} 0 & b_1 a_{1,2} & \cdots & b_1 a_{1,n} \\ b_2 a_{2,1} & 0 & \cdots & b_2 a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_n a_{n,1} & b_n a_{n,2} & \cdots & 0 \end{pmatrix} =$$

$$\begin{pmatrix} 0 & (b_2 - b_1) a_{1,2} & \cdots & (b_n - b_1) a_{1,n} \\ (b_1 - b_2) a_{2,1} & 0 & \cdots & (b_n - b_2) a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ (b_1 - b_n) a_{n,1} & (b_2 - b_n) a_{n,2} & \cdots & 0 \end{pmatrix} = ((b_j - b_i) a_{i,j}).$$

Então dada uma matriz $Y = (c_{i,j}) \in V$, tomando $Z = ((b_j - b_i)^{-1}c_{i,j})$ para $i \neq j$ e zeros na diagonal principal, temos $Y = [Z, S]$, o que prova a igualdade.

Desta forma podemos concluir que $f(S, x_2, S, S)$ consiste de todas as matrizes com apenas zeros na diagonal principal. Pela Proposição 3.4, $sl_n(\mathbb{K}) \subset f(M_n(\mathbb{K}))$. O tratamento para os casos $\alpha_2 \neq 0$ e $\alpha_3 \neq 0$ se faz de forma análoga.

Supondo $\alpha_1 = \alpha_2 = \alpha_3 = 0$, teremos

$$f(x_1, x_2, x_3, x_4) = \alpha_{1234}[x_1, x_2][x_3, x_4] + \alpha_{1324}[x_1, x_3][x_2, x_4] + \alpha_{1423}[x_1, x_4][x_2, x_3]$$
$$+\alpha_{2314}[x_2, x_3][x_1, x_4] + \alpha_{2413}[x_2, x_4][x_1, x_3] + \alpha_{3412}[x_3, x_4][x_1, x_2].$$

Consideremos dois casos.

1º caso: $\alpha_{1234} = \alpha_{2314} = \alpha_{3412} = \alpha_{1423} = -\alpha_{1324} = -\alpha_{2413}$.

Aqui temos que f é um múltiplo do polinômio standard de grau 4 e portanto basta usar a Proposição 3.15.

2º caso: Ao menos uma das igualdades $\alpha_{1234} = \alpha_{2314} = \alpha_{3412} = \alpha_{1423} = -\alpha_{1324} = -\alpha_{2413}$ não vale.

Afirmamos que existem $A, B, C \in M_n(\mathbb{K})$, tais que ao menos uma das seguintes expressões é não nula:

$$f(A, A, B, C) = (\alpha_{1324} + \alpha_{2314})[A, B][A, C] + (\alpha_{1423} + \alpha_{2413})[A, C][A, B] \quad (3.9)$$
$$f(A, B, A, C) = (\alpha_{1234} - \alpha_{2314})[A, B][A, C] + (\alpha_{3412} - \alpha_{1423})[A, C][A, B] \quad (3.10)$$
$$f(A, B, C, A) = (-\alpha_{1234} - \alpha_{2413})[A, B][A, C] + (-\alpha_{1324} - \alpha_{3412})[A, C][A, B] \quad (3.11)$$
$$f(B, A, A, C) = (-\alpha_{1234} - \alpha_{1324})[A, B][A, C] + (-\alpha_{2413} - \alpha_{3412})[A, C][A, B] \quad (3.12)$$
$$f(B, A, C, A) = (-\alpha_{1423} + \alpha_{1234})[A, B][A, C] + (\alpha_{3412} - \alpha_{2314})[A, C][A, B] \quad (3.13)$$
$$f(B, C, A, A) = (\alpha_{1324} + \alpha_{1423})[A, B][A, C] + (\alpha_{2314} + \alpha_{2413})[A, C][A, B]. \quad (3.14)$$

De fato, se λ_1 ou λ_2 são elementos não nulos em $\mathbb{K}$, então a expressão

$$\lambda_1[A, B][A, C] + \lambda_2[A, C][A, B]$$

é não nula, pois tomando por exemplo $A = e_{1,1}, B = e_{1,2}$ e $C = -e_{2,1}$, temos que ela é igual à $\lambda_1 e_{1,1} + \lambda_2 e_{2,2}$. Portanto basta analisarmos os escalares nas expressões acima. Para os casos $\alpha_{1234} \neq \alpha_{2314}, \alpha_{1234} \neq \alpha_{1423}, \alpha_{1234} \neq -\alpha_{1324}, \alpha_{1234} \neq -\alpha_{2413}, \alpha_{2314} \neq \alpha_{3412}$, $\alpha_{2314} \neq -\alpha_{1324}$, $\alpha_{2314} \neq -\alpha_{2413}, \alpha_{3412} \neq \alpha_{1423}, \alpha_{3412} \neq -\alpha_{1324}, \alpha_{3412} \neq -\alpha_{2413}$, $\alpha_{1423} \neq -\alpha_{1324}$ e $\alpha_{1423} \neq -\alpha_{2413}$, as expressões (3.10), (3.13), (3.12), (3.11), (3.13), (3.9), (3.14), (3.10), (3.11), (3.12), (3.14) e (3.9) são não nulas, respectivamente. Além disso,

- Para $\alpha_{1234} \neq \alpha_{3412}$, se $\alpha_{2314} \neq \alpha_{1234}$ então (3.10) é não nula e se $\alpha_{2314} \neq \alpha_{3412}$, então (3.13) é não nula.

- Para $\alpha_{2314} \neq \alpha_{1423}$, se $\alpha_{1234} \neq \alpha_{2314}$ então (3.10) é não nula e se $\alpha_{1234} \neq \alpha_{1423}$ então (3.13) é não nula.
- Para $-\alpha_{1324} \neq -\alpha_{2413}$, se $\alpha_{1234} \neq -\alpha_{1324}$ então (3.9) é não nula e se $\alpha_{1234} \neq -\alpha_{2413}$ então (3.11) é não nula.

Notemos que dessa forma reduzimos nosso problema ao polinômio

$$f(x_1, x_2, x_3) = [x_1, x_2][x_1, x_3] + \lambda[x_1, x_3][x_1, x_2],$$

$\lambda \in \mathbb{K}$, que já foi estudado nos Lemas 3.12 e 3.13. Com isso concluímos a demonstração do teorema. $\square$

Capítulo 4

IMAGENS DE POLINÔMIOS MULTILINEARES DE GRAU ATÉ 4 SOBRE UT_n

Neste capítulo iremos descrever o subespaço gerado pela imagem de um polinômio multilinear sobre a álgebra das matrizes triangulares superiores, bem como a imagem de polinômios multilineares de grau até quatro também sobre essa álgebra. Ressaltamos que nessa última descrição serão usadas as técnicas para reduções polinomiais apresentadas no capítulo anterior.

Os resultados deste capítulo são originais e estão no artigo [12].

4.1 O subespaço gerado pela imagem de polinômios multilineares sobre UT_n

Sejam $\mathbb{K}$ um corpo qualquer, n um inteiro positivo e $k \geqslant -1$ um inteiro. Denotaremos por $UT_n^{(k)}(\mathbb{K})$ o subespaço de $UT_n(\mathbb{K})$ gerado pelas matrizes elementares $e_{i,j}$ com $j - i > k$. Notemos que $UT_n^{(-1)}(\mathbb{K})$ denota a álgebra das matrizes triangulares superiores, enquanto que $UT_n^{(0)}(\mathbb{K})$ denota a álgebra das matrizes estritamente triangulares superiores.

Omitiremos o corpo $\mathbb{K}$ na notação $UT_n^{(k)}(\mathbb{K})$, caso não haja dúvida qual corpo está sendo considerado.

Sejam $m \geqslant 2$ um inteiro e $e_{i_1,j_1}, \ldots, e_{i_m,j_m}$ matrizes elementares triangulares superiores, isto é, $i_p \leqslant j_p$, para todo p. Sabemos que

$$e_{i_1,j_1} \cdots e_{i_m,j_m} \tag{4.1}$$

é não nulo e igual a e_{i_1,j_m} se, e somente se, $j_p = i_{p+1}$, para todo p.

Afirmamos que ao trocarmos a ordem do produto (4.1) obteremos 0 ou e_{i_1,j_m}. Para verificarmos esta afirmação, basta analisarmos o que ocorre quando trocamos o primeiro ou último termo em (4.1). Se $e_{i_l,j_l} \cdots e_{i_1,j_1} \cdots e_{i_m,j_m}$ é não nulo, então devemos ter $i_l \leqslant i_1$, que juntamente com $i_1 \leqslant i_l$, nos garante que $i_l = i_1$. Segue que $e_{i_l,j_l} \cdots e_{i_1,j_1} \cdots e_{i_m,j_m} = e_{i_1,j_m}$.

Analogamente, provamos que se $e_{i_1,j_1}\cdots e_{i_m,j_m}\cdots e_{i_l,j_l}$ é não nulo, então deve ser igual à e_{i_1,j_m}.

Pela breve discussão acima temos provado que se $f(x_1,\ldots,x_m)\in\mathbb{K}\langle X\rangle$ é multilinear, então sua imagem sobre matrizes elementares triangulares superiores sempre será 0 ou um múltiplo de uma matriz elementar triangular superior. Observamos que um resultado análogo para a álgebra das matrizes não é válido já que $[e_{1,2},e_{2,1}]=e_{1,1}-e_{2,2}$.

Vejamos alguns resultados preliminares.

Lema 4.1. *Seja $f(x_1,\ldots,x_m)\in\mathbb{K}\langle X\rangle$ multilinear. Suponha que um múltiplo não nulo de $e_{i,i+k-1}$ pertença a $f(UT_n)$, para alguns i e k. Então, $e_{1,k},e_{2,k+1},\ldots,e_{n-k+1,n}$ pertencem à $f(UT_n)$.*

Demonstração. Escrevemos $\alpha e_{i,i+k-1}=f(e_{i_1,j_1},\ldots,e_{i_m,j_m})$ para algum $\alpha\in\mathbb{K}-\{0\}$. Sem perda de generalidade assuma $i\neq 1$. Portanto,

$$\alpha e_{1,k}=f(e_{i_1-i+1,j_1-i+1},\ldots,e_{i_m-i+1,j_m-i+1}),$$

o que prova $e_{1,k}\in f(UT_n)$. De forma análoga provamos que $e_{2,k+1},\ldots,e_{n-k+1,n}\in f(UT_n)$. □

Lema 4.2. *Seja $f(x_1,\ldots,x_m)\in\mathbb{K}\langle X\rangle$ multilinear. Suponha que um múltiplo não nulo de $e_{i,i+k-1}$ pertença a $f(UT_n)$, para alguns i e k. Então, $e_{i,i+k}\in f(UT_n)$.*

Demonstração. Escrevemos $\alpha e_{i,i+k-1}=f(e_{i_1,j_1},\ldots,e_{i_m,j_m})$, para algum $\alpha\in\mathbb{K}-\{0\}$. Logo, $i+k-1=j_l$ para alguns índices $l\in\{1,\ldots,m\}$. Para cada l, substituímos o correspondente j_l por j_l+1 e, então, obtemos

$$\alpha e_{i,i+k}=f(e_{i_1,j_1},\ldots,e_{i_l,j_l+1},\ldots,e_{i_m,j_m}),$$

provando assim o lema. □

Definição 4.3. *Seja $A=\sum_{i,j=1}^{n}a_{i,j}e_{i,j}\in UT_n$. Para cada $k\in\{1,\ldots,n\}$, a k-ésima diagonal de A é aquela formada pelas entradas $(1,k),(2,k+1),\ldots,(n-k+1,n)$. Dizemos que a k-ésima diagonal de A é não nula se ao menos uma de suas entradas é não nula.*

Estamos aptos a provar o resultado principal desta seção.

Proposição 4.4. *Seja $f(x_1,\ldots,x_m)\in\mathbb{K}\langle X\rangle$ multilinear. Então, span$(f(UT_n))$ é igual a $\{0\}$ ou $UT_n^{(k)}$, para algum $k\geqslant -1$.*

Demonstração. Assuma que $f(UT_n)$ é diferente de zero. Segue que se $A\in f(UT_n)$ é não nula, então escrevendo A como combinação linear de avaliações de f em matrizes elementares triangulares superiores, obtemos que um múltiplo não nulo de $e_{i,j}$ pertence à $f(UT_n)$ para cada entrada (i,j) não nula de A.

Seja k o menor inteiro tal que a k-ésima diagonal de alguma matriz $A = \sum_{i,j=1}^{n} a_{i,j}e_{i,j} \in f(UT_n)$ é não nula. Então, existe algum $a_{i,i+k-1} \neq 0$ e, portanto, $\alpha e_{i,i+k-1} = f(e_{i_1,j_1}, \ldots, e_{i_m,j_m})$ para algum $\alpha \in \mathbb{K} - \{0\}$. Pelo Lema 4.1, as matrizes elementares $e_{1,k}, e_{2,k+1}, \ldots, e_{n-k+1,n}$ pertencem à $f(UT_n)$. Pelo Lema 4.2, também temos que $e_{i,i+k}$ pertence à $f(UT_n)$. Usando esses dois lemas alternadamente obtemos $UT_n^{(k-2)} \subset \operatorname{span}(f(UT_n))$. Logo, a igualdade $f(UT_n) = UT_n^{(k-2)}$ segue pela minimalidade de k. □

Baseados na proposição acima, conjecturamos o seguinte.

Conjectura 4.5. *A imagem de um polinômio multilinear sobre a álgebra das matrizes triangulares superiores é $\{0\}$ ou $UT_n^{(k)}$ para algum $k \geqslant -1$.*

Neste capítulo iremos provar a Conjectura 4.5 para polinômios de grau menor ou igual a 4, sob algumas condições sobre o corpo $\mathbb{K}$.

O próximo lema será de fundamental importância no restante do capítulo e afirma que nenhum subconjunto estritamente entre $UT_n^{(0)}$ e UT_n pode ser realizado como imagem de um polinômio multilinear sobre UT_n.

Lema 4.6. *Seja*

$$f(x_1, \ldots, x_m) = \sum_{\sigma \in S_m} \alpha_\sigma x_{\sigma(1)} \cdots x_{\sigma(m)}, \alpha_\sigma \in \mathbb{K} \tag{4.2}$$

um polinômio multilinear não nulo.

(i) Se $\sum_{\sigma \in S_m} \alpha_\sigma \neq 0$, então $f(UT_n) = UT_n$;

(ii) Se $\sum_{\sigma \in S_m} \alpha_\sigma = 0$ e $UT_n^{(0)} \subset f(UT_n)$, então $f(UT_n) = UT_n^{(0)}$.

Demonstração. O item (i) segue do Lema 2.17 (iv).

Para provarmos o item (ii) consideramos $\tau \in S_m$ tal que $\alpha_\tau \neq 0$. Tal permutação existe já que f é não nulo. Então, $\alpha_\tau = -\sum_{\sigma \in S_m - \{\tau\}} \alpha_\sigma$. Logo, podemos reescrever f como

$$f(x_1, \ldots, x_m) = \sum_{\sigma \in S_m - \{\tau\}} \alpha_\sigma (x_{\sigma(1)} \ldots x_{\sigma(m)} - x_{\tau(1)} \ldots x_{\tau(m)}). \tag{4.3}$$

Portanto, quando substituirmos cada variável $x_1, \ldots, x_m$ por matrizes triangulares superiores obteremos uma matriz estritamente triangular superior. De fato, a diagonal principal de um produto de matrizes triangulares superiores mantém-se a mesma, não importa a ordem em que multiplicamos. Segue que $f(UT_n) \subset UT_n^{(0)}$ e, pela hipótese, podemos concluir a igualdade desejada. $\square$

Como uma primeira aplicação do lema acima, descrevemos a imagem de um polinômio multilinear de grau arbitrário sobre UT_2. Seja $f(x_1, \dots, x_m) = \sum_{\sigma \in S_m} \alpha_\sigma x_{\sigma(1)} \cdots x_{\sigma(m)}$ multilinear não nulo. Se $\sum_{\sigma \in S_m} \alpha_\sigma \neq 0$, então $f(UT_2) = UT_2$. Caso contrário, pelo Lema 4.6 (ii), temos que $f(UT_2) \subset UT_2^{(0)}$. A partir de agora consideramos dois casos: se f é uma identidade polinomial para UT_2, então $f(UT_2) = \{0\}$ e, caso contrário, sendo a imagem de f fechada por produto por escalar, obtemos que $f(UT_2) = UT_2^{(0)}$. Assim, obtemos o seguinte resultado, que confirma a validade da Conjectura 4.5 para $n = 2$ e polinômios de grau arbitrário.

Proposição 4.7. *Seja $\mathbb{K}$ um corpo qualquer. A imagem de um polinômio $f \in \mathbb{K}\langle X\rangle$ multilinear de grau arbitrário sobre UT_2 é $\{0\}, UT_2^{(0)}$ ou UT_2.*

4.2 Polinômios multilineares de grau dois sobre UT_n

O objetivo desta seção é a descrição das imagens de polinômios multilineares de grau dois sobre UT_n.

Teorema 4.8. *Sejam $\mathbb{K}$ um corpo qualquer e $f(x_1, x_2) \in \mathbb{K}\langle X\rangle$ multilinear de grau dois. Então, $f(UT_n)$ é $\{0\}, UT_n^{(0)}$ ou UT_n.*

Demonstração. Escrevemos $f(x_1, x_2) = \alpha x_1 x_2 + \beta x_2 x_1, \alpha, \beta \in \mathbb{K}$.

Caso $\alpha + \beta \neq 0$, então pelo Lema 4.6 (i) temos $f(UT_n) = UT_n$. Assim podemos supor que $\alpha + \beta = 0$. Se $\alpha = \beta = 0$, então $f(UT_n) = \{0\}$. Caso contrário, podemos assumir $f(x_1, x_2) = x_1 x_2 - x_2 x_1$.

Sejam $A = \sum_{i=1}^{n-1} e_{i,i+1}$ e $B = \sum_{i,j=1}^{n} b_{i,j} e_{i,j} \in UT_n$. Logo,

$$AB - BA = \sum_{i=1}^{n-1} \sum_{j=2}^{n} (b_{i+1,j} - b_{i,j-1}) e_{i,j}. \tag{4.4}$$

Para $C = \sum_{i,j=1}^{n} c_{i,j} e_{i,j} \in UT_n^{(0)}$, consideramos o seguinte sistema linear

$$\begin{cases} b_{2,2} - b_{1,1} & = & c_{1,2} \\ & \vdots & \\ b_{2,n} - b_{1,n-1} & = & c_{1,n} \\ & \vdots & \\ b_{n,n} - b_{n-1,n-1} & = & c_{n-1,n} \end{cases} \tag{4.5}$$

Uma solução deste sistema é $b_{1,j} = 0$ para $j \in \{1, \ldots, n\}$ e $b_{i+1,j} = c_{i,j} + c_{i-1,j-1} + \cdots + c_{1,j-(i-1)}$ em que $i < j$, $i \in \{2, \ldots, n-1\}$ e $j \in \{2, \ldots, n\}$. Portanto, $UT_n^{(0)} \subset f(UT_n)$ e pelo Lema 4.6 (ii) obtemos $f(UT_n) = UT_n^{(0)}$. □

4.3 Polinômios multilineares de grau três sobre UT_n

Nesta seção iremos descrever as imagens de polinômios multilineares de grau três sobre UT_n. Antes, vejamos um lema técnico.

Lema 4.9. *Sejam $\mathbb{K}$ um corpo com pelo menos n elementos e $d_{1,1}, \ldots, d_{n,n} \in \mathbb{K}$ elementos distintos entre si. Então para $D = diag(d_{1,1}, \ldots, d_{n,n})$ e $k \geqslant 0$ temos*

$$[UT_n^{(k)}, D] = UT_n^{(k)} \text{ e } [UT_n, D] = UT_n^{(0)}.$$

Demonstração. Obviamente $[UT_n^{(k)}, D] \subset UT_n^{(k)}$.

Seja $A = \sum_{j-i>k} a_{i,j} e_{i,j}$ um elemento qualquer de $UT_n^{(k)}$. Então,

$$AD - DA = \sum_{j-i>k} a_{i,j}(d_{j,j} - d_{i,i}) e_{i,j}. \tag{4.6}$$

Portanto, se $B = \sum_{j-i>k} b_{i,j} e_{i,j} \in UT_n^{(k)}$, então escolhemos $a_{i,j} = b_{i,j}(d_{j,j} - d_{i,i})^{-1}$ para $j - i > k$. Isto prova que $[UT_n^{(k)}, D] \supset UT_n^{(k)}$, provando assim a primeira igualdade.

Provemos a segunda igualdade. É imediato que $[UT_n, D] \subset UT_n^{(0)}$. Desde que $UT_n^{(0)} \subset UT_n$, segue que $[UT_n^{(0)}, D] \subset [UT_n, D]$. Pela primeira equação deste lema para $k = 0$, obtemos $UT_n^{(0)} = [UT_n^{(0)}, D] \subset [UT_n, D]$, o que prova a segunda igualdade. □

A seguir é apresentado o resultado principal desta seção.

Teorema 4.10. *Sejam $\mathbb{K}$ um corpo com pelo menos n elementos e $f(x_1, x_2, x_3) \in \mathbb{K}\langle X \rangle$ multilinear. Então, $f(UT_n)$ é $\{0\}, UT_n^{(0)}$ ou UT_n.*

Demonstração. Suponha que f é não nulo. Se $f(1, x_2, x_3) \neq 0, f(x_1, 1, x_3) \neq 0$ ou $f(x_1, x_2, 1) \neq 0$ então pelo Teorema 4.8 obtemos $UT_n^{(0)} \subset f(UT_n)$. Pelo Lema 4.6, temos $f(UT_n) = UT_n^{(0)}$ ou $f(UT_n) = UT_n$.

Portanto, podemos assumir $f(1, x_2, x_3) = f(x_1, 1, x_3) = f(x_1, x_2, 1) = 0$. Segue que f é um polinômio multilinear próprio de grau 3, e portanto é um polinômio multilinear de Lie de grau 3. Pelo Exemplo 1.43, podemos escrever f como

$$f(x_1, x_2, x_3) = \alpha_1[[x_2, x_1], x_3] + \alpha_2[[x_3, x_1], x_2], \alpha_1, \alpha_2 \in \mathbb{K},$$

e sem perda de generalidade podemos supor

$$f(x_1, x_2, x_3) = [[x_2, x_1], x_3] + \alpha[[x_3, x_1], x_2], \alpha \in \mathbb{K}.$$

Pelo Lema 4.9, $UT_n^{(0)} = [UT_n^{(0)}, D] = [[UT_n, D], D]$. Logo, substituindo x_1 e x_3 por D e x_2 por matrizes em UT_n obtemos $f(UT_n) = UT_n^{(0)}$. □

4.4 Polinômios multilineares de grau quatro sobre UT_n

Nesta seção iremos descrever as imagens de polinômios multilineares de grau quatro sobre UT_n. Comecemos com o seguinte lema.

Lema 4.11. *Seja $\mathbb{K}$ um corpo qualquer. Então, $[UT_n^{(0)}, UT_n^{(0)}] = UT_n^{(1)}$.*

Demonstração. Certamente, $[UT_n^{(0)}, UT_n^{(0)}] \subset UT_n^{(1)}$. Considere $A = \sum_{i=1}^{n-1} e_{i,i+1}$ e $B = \sum_{i,j=1}^{n} b_{i,j}e_{i,j} \in UT_n^{(0)}$. Segue que

$$[A, B] = \sum_{i=1}^{n-1}\sum_{j=2}^{n}(b_{i+1,j} - b_{i,j-1})e_{i,j}.$$

Portanto, dada uma matriz $C = \sum_{i,j=1}^{n} c_{i,j}e_{i,j} \in UT_n^{(1)}$, basta encontrarmos uma solução do seguinte sistema

$$\begin{cases} b_{2,3} - b_{1,2} & = & c_{1,3} \\ & \vdots & \\ b_{2,n} - b_{1,n-1} & = & c_{1,n} \\ & \vdots & \\ b_{n-1,n} - b_{n-2,n-1} & = & c_{n-2,n} \end{cases} \tag{4.7}$$

E, de fato, uma solução de (4.7) é dada tomando $b_{1,i} = 0$ para $i \in \{2, \dots, n-1\}$ e $b_{i+1,j} = c_{i,j} + \cdots + c_{1,j-(i-1)}$ para $i \in \{1, \dots, n-2\}, j \in \{3, \dots, n\}$. □

Teorema 4.12. *Sejam $\mathbb{K}$ um corpo de característica zero e $f(x_1, x_2, x_3, x_4) \in \mathbb{K}\langle X\rangle$ multilinear. Então, $f(UT_n)$ é $\{0\}, UT_n^{(1)}, UT_n^{(0)}$ ou UT_n.*

Demonstração. Assuma que f é um polinômio não nulo. Se

$$f(1, x_2, x_3, x_4), f(x_1, 1, x_3, x_4), f(x_1, x_2, 1, x_4) \text{ ou } f(x_1, x_2, x_3, 1)$$

é não nulo, então pelo Lema 4.6 e pelo Teorema 4.10 obtemos que $f(UT_n)$ é $UT_n^{(0)}$ ou UT_n.

Portanto, podemos supor

$$f(1, x_2, x_3, x_4) = f(x_1, 1, x_3, x_4) = f(x_1, x_2, 1, x_4) = f(x_1, x_2, x_3, 1) = 0.$$

Pelo Corolário 1.47 podemos escrever f como

$$\begin{aligned} f(x_1, x_2, x_3, x_4) &= L(x_1, x_2, x_3, x_4) + \alpha_{1234}[x_1, x_2][x_3, x_4] + \alpha_{1324}[x_1, x_3][x_2, x_4] \\ &+ \alpha_{1423}[x_1, x_4][x_2, x_3] + \alpha_{2314}[x_2, x_3][x_1, x_4] + \alpha_{2413}[x_2, x_4][x_1, x_3] \\ &+ \alpha_{3412}[x_3, x_4][x_1, x_2], \end{aligned}$$

em que $L(x_1, x_2, x_3, x_4)$ é um polinômio de Lie e $\alpha_{1234}, \alpha_{1324}, \alpha_{1423}, \alpha_{2314}, \alpha_{2413}, \alpha_{3412} \in \mathbb{K}$.

Pelo Exemplo 1.44 podemos escrever L como

$$\begin{aligned} L(x_1, x_2, x_3, x_4) &= \alpha_1[[[x_2, x_1], x_3], x_4] + \alpha_2[[[x_3, x_1], x_2], x_4] + \alpha_3[[[x_4, x_1], x_2], x_3] \\ &+ \alpha_4[[x_4, x_1], [x_3, x_2]] + \alpha_5[[x_4, x_2], [x_3, x_1]] + \alpha_6[[x_4, x_3], [x_2, x_1]], \end{aligned}$$

em que $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6 \in \mathbb{K}$.

Desenvolvendo os colchetes nos três últimos termos de L, podemos assumir que f se escreve como

$$\begin{aligned} f(x_1, x_2, x_3, x_4) &= \alpha_1[[[x_2, x_1], x_3], x_4] + \alpha_2[[[x_3, x_1], x_2], x_4] + \alpha_3[[[x_4, x_1], x_2], x_3] \\ &+ \alpha_{1234}[x_1, x_2][x_3, x_4] + \alpha_{1324}[x_1, x_3][x_2, x_4] + \alpha_{1423}[x_1, x_4][x_2, x_3] \\ &+ \alpha_{2314}[x_2, x_3][x_1, x_4] + \alpha_{2413}[x_2, x_4][x_1, x_3] + \alpha_{3412}[x_3, x_4][x_1, x_2]. \end{aligned}$$

Digamos que para algum $i \in \{1, 2, 3\}$ tenhamos $\alpha_i \neq 0$, digamos sem perda de generalidade que seja $\alpha_1 \neq 0$. Portanto, substituindo x_1, x_3 e x_4 por uma matriz $D = diag(d_{1,1}, \ldots, d_{n,n})$ em que $d_{1,1}, \ldots, d_{n,n}$ são elementos distintos entre si em $\mathbb{K}$, obtemos $f(D, x_2, D, D) = \alpha_1[[[x_2, D], D], D]$. Segue que $f(UT_n) = UT_n^{(0)}$, pelo Lema 4.9. Dessa forma, iremos supor que $\alpha_1 = \alpha_2 = \alpha_3 = 0$ e, portanto, reescrever f como

$$\begin{aligned} f(x_1, x_2, x_3, x_4) &= \alpha_{1234}[x_1, x_2][x_3, x_4] + \alpha_{1324}[x_1, x_3][x_2, x_4] + \alpha_{1423}[x_1, x_4][x_2, x_3] \\ &+ \alpha_{2314}[x_2, x_3][x_1, x_4] + \alpha_{2413}[x_2, x_4][x_1, x_3] + \alpha_{3412}[x_3, x_4][x_1, x_2]. \end{aligned}$$

Observamos que podemos supor $n \geqslant 3$, já que para $n = 2$, o polinômio f acima é uma identidade polinomial para UT_2 (Exemplo 1.34). Neste caso, claramente $f(UT_n) \subset UT_n^{(1)}$. Agora nosso propósito é provar $f(UT_n) \supset UT_n^{(1)}$ nos dois casos seguintes.

Caso 1. Assuma $\alpha_{1234} = \alpha_{2314} = \alpha_{3412} = \alpha_{1423} = -\alpha_{1324} = -\alpha_{2413}$. Então, podemos supor

$$\begin{aligned} f(x_1, x_2, x_3, x_4) &= [x_1, x_2][x_3, x_4] + [x_3, x_4][x_1, x_2] + [x_2, x_3][x_1, x_4] \\ &- [x_1, x_3][x_2, x_4] - [x_2, x_4][x_1, x_3]. \end{aligned}$$

Sejam $A \in UT_n^{(1)}$ e $D = diag(d_{1,1}, \ldots, d_{n,n})$, em que $d_{1,1}, \ldots, d_{n,n}$ são elementos distintos entre si em $\mathbb{K}$. Então pelo Lema 4.9 existe $G \in UT_n^{(1)}$ tal que $A = [D, G]$. Pelo Lema 4.11 existem $E, F \in UT_n^{(0)}$ tais que $G = [E, F]$. Novamente pelo Lema 4.9 existem $B, C \in UT_n$ tais que $E = [D, B]$ e $F = [D, C]$. Segue que $A = [D, [[D, B], [D, C]]]$. Observando que

$$\begin{aligned} f(D, D^2, B, C) &= [D^2, B][D, C] + [D, C][D^2, B] - [D, B][D^2, C] - [D^2, C][D, B] \\ &= [D, [[D, B], [D, C]]], \end{aligned}$$

podemos afirmar que $A \in UT_n^{(1)}$, provando assim que $f(UT_n) = UT_n^{(1)}$.

Caso 2. Assuma que ao menos uma das igualdades

$$\alpha_{1234} = \alpha_{2314} = \alpha_{3412} = \alpha_{1423} = -\alpha_{1324} = -\alpha_{2413}$$

não ocorra. Segue que existem $A, B, C \in UT_n$ tais que ao menos uma das matrizes abaixo é não nula

$$\begin{aligned} f(A, A, B, C) &= (\alpha_{1324} + \alpha_{2314})[A, B][A, C] + (\alpha_{1423} + \alpha_{2413})[A, C][A, B], \\ f(A, B, A, C) &= (\alpha_{1234} + \alpha_{2314})[A, B][A, C] + (\alpha_{3412} + \alpha_{1423})[A, C][A, B], \\ f(A, B, C, A) &= (-\alpha_{1234} - \alpha_{2413})[A, B][A, C] + (-\alpha_{1324} - \alpha_{3412})[A, C][A, B], \\ f(B, A, A, C) &= (-\alpha_{1234} - \alpha_{1324})[A, B][A, C] + (-\alpha_{2413} - \alpha_{3412})[A, C][A, B], \\ f(B, A, C, A) &= (-\alpha_{1423} + \alpha_{1234})[A, B][A, C] + (\alpha_{3412} - \alpha_{2314})[A, C][A, B], \\ f(B, C, A, A) &= (\alpha_{1324} + \alpha_{1423})[A, B][A, C] + (\alpha_{2314} + \alpha_{2413})[A, C][A, B]. \end{aligned}$$

De fato, se $\lambda_1, \lambda_2 \in \mathbb{K}$, em que sem perda de generalidade λ_1 é não nulo, então a expressão

$$\lambda_1[A, B][A, C] + \lambda_2[A, C][A, B] \tag{4.8}$$

é não nula, pois podemos tomar $A = e_{2,2}, B = e_{1,2}$ e $C = e_{2,3}$. Logo é suficiente apenas analisar os escalares que aparecem nas expressões acima de forma análoga à demonstração do Teorema 3.16.

Portanto, basta provarmos que para $\lambda \in \mathbb{K}$ arbitrário, qualquer matriz em $UT_n^{(1)}$ pode ser escrita como

$$[A, B][A, C] + \lambda[A, C][A, B],$$

para algumas matrizes $A, B, C \in UT_n$. Usando o Lema 4.9 e tomando $A = diag(a_{1,1}, \dots, a_{n,n})$ em que $a_{1,1}, \dots, a_{n,n}$ são elementos distintos entre si em $\mathbb{K}$, existe $B \in UT_n$ tal que $\sum_{i=1}^{n-1} e_{i,i+1} = [A, B]$. Escrevendo $[A, C] = \sum_{i,j=1}^{n} b_{i,j} e_{i,j}$ temos

$$[A, B][A, C] + \lambda[A, C][A, B] = \sum_{i=1}^{n-1} \sum_{j=2}^{n} (b_{i+1,j} - \lambda b_{i,j-1}) e_{i,j}.$$

Logo, para $M = \sum_{i,j=1}^{n} c_{i,j} e_{i,j} \in UT_n^{(1)}$, o seguinte sistema possui solução.

$$\begin{cases} b_{2,3} + \lambda b_{1,2} & = & c_{1,3} \\ & \vdots & \\ b_{2,n} + \lambda b_{1,n-1} & = & c_{1,n} \\ & \vdots & \\ b_{n-1,n} + \lambda b_{n-2,n-1} & = & c_{n-2,n} \end{cases}$$

De fato, basta tomarmos $b_{1,j} = 0$ para $j \in \{2, \dots, n-1\}$ e

$$b_{i+1,j} = c_{i,j} - \lambda c_{i-1,j-1} + \cdots + (-\lambda)^{i-1} c_{1,j-(i-1)},$$

para $i \in \{1, \dots, n-2\}, j \in \{3, \dots, n\}$. Portanto, $M \in f(UT_n)$, encerrando assim a prova do teorema. $\square$

Capítulo 5

IMAGENS DE POLINÔMIOS MULTILINEARES SOBRE $UT_n^{(0)}$

Neste último capítulo iremos descrever as imagens de polinômios multilineares de grau arbitrário sobre a álgebra das matrizes estritamente triangulares superiores. Os resultados que serão expostos são originais e podem ser encontrados em [11].

Durante todo o capítulo, $\mathbb{K}$ denotará um corpo qualquer.

5.1 Um lema técnico

Iniciamos esta seção com algumas notações.

Sejam $Y = \{y_k^{(l)} | k, l \in \{1, \ldots, n\}\}$ um conjunto de variáveis comutativas e $\mathbb{K}[Y]$ a álgebra comutativa livre, livremente gerada pelo conjunto Y.

Relembrando que S_m denota o grupo das permutações dos elementos do conjunto $\{1, \ldots, m\}$, para cada $j \in \{1, \ldots, m\}$ denotaremos:

$$S_m^{(j)} = \{\sigma \in S_m | \sigma(j) = j\} \text{ e } G^{(j)} = S_m^{(1)} \cap S_m^{(j)} \cap S_m^{(j+1)} \cap \cdots \cap S_m^{(m)}.$$

Também denotaremos $S_m^{(1)}$ por $G^{(m+1)}$.

O principal objetivo desta seção é provar o seguinte lema.

Lema 5.1. *Sejam $\sigma \in S_m^{(1)}$ e $\alpha_\sigma \in \mathbb{K}$, em que $\alpha_{id} = 1$. Então, podemos substituir as variáveis $y_2^{(2)}, \ldots, y_{n-1}^{(2)}, \ldots, y_2^{(m)}, \ldots, y_{n-1}^{(m)}$ por escalares em $\mathbb{K}$ tais que as seguintes* *expressões são não nulas*

$$\begin{cases} \sum\limits_{\sigma \in S_m^{(1)}} \alpha_\sigma y_2^{(\sigma(2))} \cdots y_m^{(\sigma(m))} \\ \sum\limits_{\sigma \in S_m^{(1)}} \alpha_\sigma y_3^{(\sigma(2))} \cdots y_{m+1}^{(\sigma(m))} \\ \vdots \\ \sum\limits_{\sigma \in S_m^{(1)}} \alpha_\sigma y_{n-m+1}^{(\sigma(2))} \cdots y_{n-1}^{(\sigma(m))} \end{cases} \tag{5.1}$$

Apresentamos o Lema 5.1 para valores pequenos de m antes de fornecermos uma prova geral, a fim de que o leitor possa compreender melhor a demonstração do caso geral.

Observamos que o caso $m = 2$ do Lema 5.1 é trivial.

Lema 5.2. *Sejam $\sigma \in S_3^{(1)}$ e $\alpha_\sigma \in \mathbb{K}$, em que $\alpha_{id} = 1$. Então, podemos substituir as variáveis $y_2^{(2)}, \ldots, y_{n-1}^{(2)}, y_2^{(3)}, \ldots, y_{n-1}^{(3)}$ por escalares em $\mathbb{K}$ tais que as seguintes expressões são não nulas*

$$\begin{cases} \sum\limits_{\sigma \in S_3^{(1)}} \alpha_\sigma y_2^{(\sigma(2))} y_3^{(\sigma(3))} \\ \sum\limits_{\sigma \in S_3^{(1)}} \alpha_\sigma y_3^{(\sigma(2))} y_4^{(\sigma(3))} \\ \vdots \\ \sum\limits_{\sigma \in S_3^{(1)}} \alpha_\sigma y_{n-2}^{(\sigma(2))} y_{n-1}^{(\sigma(3))} \end{cases} \tag{5.2}$$

Demonstração. Desde que $S_3^{(1)} = \{id, (23)\}$, então podemos reescrever (5.2) como

$$\begin{cases} y_2^{(2)} y_3^{(3)} + \alpha_{(23)} y_2^{(3)} y_3^{(2)} \\ y_3^{(2)} y_4^{(3)} + \alpha_{(23)} y_3^{(3)} y_4^{(2)} \\ \vdots \\ y_{n-2}^{(2)} y_{n-1}^{(3)} + \alpha_{(23)} y_{n-2}^{(3)} y_{n-1}^{(2)} \end{cases} \tag{5.3}$$

Se $\alpha_{(23)} = 0$, então basta substituir todas as variáveis em (5.3) por 1.

Se $\alpha_{(23)} \neq 0$, então fazemos a seguinte substituição

$$\begin{cases} y_k^{(2)} \text{ por } 0 \text{ e } y_k^{(3)} \text{ por } 1; \text{ se } k \text{ é ímpar} \\ y_k^{(2)} \text{ por } 1 \text{ e } y_k^{(3)} \text{ por } 0; \text{ se } k \text{ é par,} \end{cases}$$

em que $k \in \{2, \ldots, n-1\}$. Dessa forma, cada avaliação em (5.3) será 1 ou $\alpha_{(23)}$. $\square$

Antes da prova geral do Lema 5.1 para $m \geqslant 4$, iremos ilustrá-la no caso $m = 4$.

Lema 5.3. *Seja $\sigma \in S_4^{(1)}$ e $\alpha_\sigma \in \mathbb{K}$ em que $\alpha_{id} = 1$. Então podemos substituir as variáveis $y_2^{(2)}, \ldots, y_{n-1}^{(2)}, y_2^{(3)}, \ldots, y_{n-1}^{(3)}, y_2^{(4)}, \ldots, y_{n-1}^{(4)}$ por escalares em $\mathbb{K}$ tais que as seguintes expressões são não nulas*

$$\begin{cases} \sum\limits_{\sigma\in S_4^{(1)}} \alpha_\sigma y_2^{(\sigma(2))} y_3^{(\sigma(3))} y_4^{(\sigma(4))} \\ \sum\limits_{\sigma\in S_4^{(1)}} \alpha_\sigma y_3^{(\sigma(2))} y_4^{(\sigma(3))} y_5^{(\sigma(4))} \\ \qquad\qquad\vdots \\ \sum\limits_{\sigma\in S_4^{(1)}} \alpha_\sigma y_{n-3}^{(\sigma(2))} y_{n-2}^{(\sigma(3))} y_{n-1}^{(\sigma(4))} \end{cases}$$

Demonstração. Inicialmente, reescrevemos o sistema acima como

$$\begin{cases} \left(\sum\limits_{\sigma\in G^{(4)}} \alpha_\sigma y_2^{(\sigma(2))} y_3^{(\sigma(3))}\right) y_4^{(4)} + \sum\limits_{\sigma\in S_4^{(1)}-G^{(4)}} \alpha_\sigma y_2^{(\sigma(2))} y_3^{(\sigma(3))} y_4^{(\sigma(4))} \\ \left(\sum\limits_{\sigma\in G^{(4)}} \alpha_\sigma y_3^{(\sigma(2))} y_4^{(\sigma(3))}\right) y_5^{(4)} + \sum\limits_{\sigma\in S_4^{(1)}-G^{(4)}} \alpha_\sigma y_3^{(\sigma(2))} y_4^{(\sigma(3))} y_5^{(\sigma(4))} \\ \qquad\qquad\qquad\qquad\vdots \\ \left(\sum\limits_{\sigma\in G^{(4)}} \alpha_\sigma y_{n-3}^{(\sigma(2))} y_{n-2}^{(\sigma(3))}\right) y_{n-1}^{(4)} + \sum\limits_{\sigma\in S_4^{(1)}-G^{(4)}} \alpha_\sigma y_{n-3}^{(\sigma(2))} y_{n-2}^{(\sigma(3))} y_{n-1}^{(\sigma(4))} \end{cases} \tag{5.4}$$

e então a prova será obtida pelos próximos dois passos

Passo 1: Afirmamos que, para uma substituição adequada de variáveis, os polinômios

$$\sum_{\sigma\in G^{(4)}} \alpha_\sigma y_{k+1}^{(\sigma(2))} y_{k+2}^{(\sigma(3))} \tag{5.5}$$

assumem valores não nulos em $\mathbb{K}$, para todo $k \in \{1, \ldots, n-4\}$. De fato, desde que $G^{(4)} = \{id, (23)\}$, usando as mesmas ideias da demonstração do Lema 5.2 podemos substituir as variáveis $y_{k+1}^{(\sigma(2))}, y_{k+2}^{(\sigma(3))}$ por escalares $\lambda_{k+1}^{(\sigma(2))}, \lambda_{k+2}^{(\sigma(3))}$ em $\mathbb{K}$ tais que $\sum\limits_{\sigma\in G^{(4)}} \alpha_\sigma \lambda_{k+1}^{(\sigma(2))} \lambda_{k+2}^{(\sigma(3))} \neq 0$, para todo $k \in \{1, \ldots, n-4\}$.

Passo 2: Procederemos aplicando um processo iterativo nos seguintes casos.

Caso 1*:* No primeiro polinômio de (5.4), para $\sigma \in G^{(4)}$, substituímos as variáveis $y_2^{(\sigma(2))}, y_3^{(\sigma(3))}$ pelos escalares $\lambda_2^{(\sigma(2))}, \lambda_3^{(\sigma(3))}$ definidos no Passo 1, respectivamente. Agora consideramos todas as variáveis $y_p^{(q)}$ dentre $y_2^{(\sigma(2))}, y_3^{(\sigma(3))}, y_4^{(\sigma(4))}$, com $\sigma \in S_4^{(1)} - G^{(4)}$. Se $y_p^{(q)}$ ocorreu no Passo 1, então apenas substituímos essa variável pelo escalar $\lambda_p^{(q)}$ definido anteriormente. Caso contrário, substituímos $y_p^{(q)}$ por qualquer escalar $\lambda_p^{(q)}$ em $\mathbb{K}$. Dessa forma, obteremos uma função linear em termos de $y_4^{(4)}$:

$$\left(\sum_{\sigma\in G^{(4)}} \alpha_\sigma \lambda_2^{(\sigma(2))} \lambda_3^{(\sigma(3))}\right) y_4^{(4)} + \sum_{\sigma\in S_4^{(1)}-G^{(4)}} \alpha_\sigma \lambda_2^{(\sigma(2))} \lambda_3^{(\sigma(3))} \lambda_4^{(\sigma(4))}. \tag{5.6}$$

Pelo Passo 1, o coeficiente de $y_4^{(4)}$ acima é não nulo. Então, podemos substituir $y_4^{(4)}$ por algum elemento $\lambda_4^{(4)}$ em $\mathbb{K}$ de sorte que (5.6) seja não nulo em $\mathbb{K}$.

Caso 2: No segundo polinômio de (5.4), para $\sigma \in G^{(4)}$, substituímos as variáveis $y_3^{(\sigma(2))}, y_4^{(\sigma(3))}$ pelos escalares $\lambda_3^{(\sigma(2))}, \lambda_4^{(\sigma(3))}$ definidos no Passo 1, respectivamente. Agora consideramos todas as variáveis $y_p^{(q)}$ dentre $y_3^{(\sigma(2))}, y_4^{(\sigma(3))}, y_5^{(\sigma(4))}$, com $\sigma \in S_4^{(1)} - G^{(4)}$. Se $y_p^{(q)}$ ocorreu no Passo 1 ou no Caso 1, então apenas substituímos essa variável pelo escalar $\lambda_p^{(q)}$ definido anteriormente. Caso contrário, substituímos $y_p^{(q)}$ por qualquer escalar $\lambda_p^{(q)}$ em $\mathbb{K}$. Dessa forma, obteremos em uma função linear em termos de $y_5^{(4)}$:

$$\left(\sum_{\sigma \in G^{(4)}} \alpha_\sigma \lambda_3^{(\sigma(2))} \lambda_4^{(\sigma(3))}\right) y_5^{(4)} + \sum_{\sigma \in S_4^{(1)} - G^{(4)}} \alpha_\sigma \lambda_3^{(\sigma(2))} \lambda_4^{(\sigma(3))} \lambda_5^{(\sigma(4))}. \tag{5.7}$$

Pelo Passo 1, o coeficiente de $y_5^{(4)}$ acima é não nulo. Então, podemos substituir $y_5^{(4)}$ por algum elemento $\lambda_5^{(4)}$ em $\mathbb{K}$ de sorte que (5.7) seja não nulo em $\mathbb{K}$.

Seguindo essas ideias, o último caso será:

Caso $n-4$: No último polinômio de (5.4), para $\sigma \in G^{(4)}$, substituímos as variáveis $y_{n-3}^{(\sigma(2))}, y_{n-2}^{(\sigma(3))}$ pelos escalares $\lambda_{n-3}^{(\sigma(2))}, \lambda_{n-2}^{(\sigma(3))}$ definidos no Passo 1, respectivamente. Agora consideramos todas as variáveis $y_p^{(q)}$ dentre $y_{n-3}^{(\sigma(2))}, y_{n-2}^{(\sigma(3))}, y_{n-1}^{(\sigma(4))}$, com $\sigma \in S_4^{(1)} - G^{(4)}$. Se $y_p^{(q)}$ ocorreu no Passo 1 ou nos casos anteriores, então apenas substituímos essa variável pelo escalar $\lambda_p^{(q)}$ definido anteriormente. Caso contrário, substituímos $y_p^{(q)}$ por qualquer escalar $\lambda_p^{(q)}$ em $\mathbb{K}$. Dessa forma, obteremos em uma função linear em termos de $y_{n-1}^{(4)}$:

$$\left(\sum_{\sigma \in G^{(4)}} \alpha_\sigma \lambda_{n-3}^{(\sigma(2))} \lambda_{n-2}^{(\sigma(3))}\right) y_{n-1}^{(4)} + \sum_{\sigma \in S_4^{(1)} - G^{(4)}} \alpha_\sigma \lambda_{n-3}^{(\sigma(2))} \lambda_{n-2}^{(\sigma(3))} \lambda_{n-1}^{(\sigma(4))}. \tag{5.8}$$

Pelo Passo 1, o coeficiente de $y_{n-1}^{(4)}$ acima é não nulo. Então, podemos substituir $y_{n-1}^{(4)}$ por algum elemento $\lambda_{n-1}^{(4)}$ em $\mathbb{K}$ de sorte que (5.8) seja não nulo em $\mathbb{K}$. Portanto, o Passo 2 conclui a demonstração do lema. $\square$

Usando as notações introduzidas no início desta seção, podemos reescrever cada polinômio

$$g_k = \sum_{\sigma \in S_m^{(1)}} \alpha_\sigma y_{k+1}^{(\sigma(2))} \cdots y_{k+m-1}^{(\sigma(m))}, k \in \{1, \ldots, n-m\}$$

do sistema (5.1) como

$$\begin{aligned} g_k = \Bigg(\Bigg(\cdots\Bigg(\Bigg(\Bigg(\sum_{\sigma \in G^{(4)}} \alpha_\sigma y_{k+1}^{(\sigma(2))} y_{k+3}^{(\sigma(3))}\Bigg) y_{k+3}^{(4)} + \sum_{\sigma \in G^{(5)} - S_m^{(4)}} \alpha_\sigma y_{k+1}^{(\sigma(2))} y_{k+2}^{(\sigma(3))} y_{k+3}^{(\sigma(4))}\Bigg) y_{k+4}^{(5)} \\ + \sum_{\sigma \in G^{(6)} - S_m^{(5)}} \alpha_\sigma y_{k+1}^{(\sigma(2))} y_{k+2}^{(\sigma(3))} y_{k+3}^{(\sigma(4))} y_{k+4}^{(\sigma(5))}\Bigg) y_{k+5}^{(6)} + \cdots\Bigg) y_{k+m-2}^{(m-1)} \\ + \sum_{\sigma \in G^{(m)} - S_m^{(m-1)}} \alpha_\sigma y_{k+1}^{(\sigma(2))} \cdots y_{k+m-2}^{(\sigma(m-1))}\Bigg) y_{k+m-1}^{(m)} + \sum_{\sigma \in S_m^{(1)} - S_m^{(m)}} \alpha_\sigma y_{k+1}^{(\sigma(2))} \cdots y_{k+m-1}^{(\sigma(m))}. \end{aligned}$$

Agora estamos aptos à provar o Lema 5.1.

Demonstração. (Lema 5.1) A prova deste lema será feita em $m-2$ passos. O primeiro é um caso especial, enquanto que para cada $j \in \{2, \ldots, m-2\}$, o j-ésimo passo consistirá em usar os passos precedentes para concluir que os polinômios

$$\left(\cdots \left(\left(\sum_{\sigma \in G^{(4)}} \alpha_\sigma y_{k+1}^{(\sigma(2))} y_{k+2}^{(\sigma(3))} \right) y_{k+3}^{(4)} + \sum_{\sigma \in G^{(5)} - S_m^{(4)}} \alpha_\sigma y_{k+1}^{(\sigma(2))} y_{k+2}^{(\sigma(3))} y_{k+3}^{(\sigma(4))} \right) y_{k+4}^{(5)} + \cdots \right) y_{j+k+1}^{(j+2)} + \sum_{\sigma \in G^{(j+3)} - S_m^{(j+2)}} \alpha_\sigma y_{k+1}^{(\sigma(2))} \cdots y_{j+k+1}^{(\sigma(j+2))} \quad (5.9)$$

assumam valores não nulos em $\mathbb{K}$, para todo $k \in \{1, \ldots, n-m\}$. Portanto, levando em consideração a escrita de cada polinômio g_k, o lema estará provado no final do último passo.

Passo 1: Afirmamos que para uma substituição adequada de variáveis, os polinômios

$$\sum_{\sigma \in G^{(4)}} \alpha_\sigma y_{k+1}^{(\sigma(2))} y_{k+2}^{(\sigma(3))}$$

assumem valores não nulos em $\mathbb{K}$, para todo $k \in \{1, \ldots, n-m\}$. De fato, desde que $G^{(4)} = \{id, (23)\}$, usando as mesmas ideias da demonstração do Lema 5.2 podemos substituir as variáveis $y_{k+1}^{(\sigma(2))}, y_{k+2}^{(\sigma(3))}$ por escalares $\lambda_{k+1}^{(\sigma(2))}, \lambda_{k+2}^{(\sigma(3))}$ em $\mathbb{K}$ tais que

$$\sum_{\sigma \in G^{(4)}} \alpha_\sigma \lambda_{k+1}^{(\sigma(2))} \lambda_{k+2}^{(\sigma(3))} \neq 0$$

para todo $k \in \{1, \ldots, n-m\}$.

Isto encerra o Passo 1. Agora, assuma que o Passo $j-1$ esteja feito, isto é, os escalares $\alpha_{k+1}^{(\sigma(2))}, \ldots, \alpha_{k+j}^{(\sigma(j+1))}$ foram definidos para $\sigma \in G^{(j+2)}$ e $k \in \{1, \ldots, n-m\}$, tais que o coeficiente de $y_{j+k+1}^{(j+2)}$ em (5.9) seja não nulo:

Passo j: Procederemos aplicando um processo iterativo nos seguintes casos.

Caso 1*:* Para $k = 1$ em (5.9) e para $\sigma \in G^{(j+2)}$, substituímos as variáveis $y_2^{(\sigma(2))}, \cdots, y_{j+1}^{(\sigma(j+1))}$ pelos escalares $\lambda_2^{(\sigma(2))}, \cdots, \lambda_{j+1}^{(\sigma(j+1))}$ definidos no Passo $j-1$, respectivamente. Agora consideramos todas as variáveis $y_p^{(q)}$ dentre $y_2^{(\sigma(2))}, \cdots, y_{j+2}^{(\sigma(j+2))}$, com $\sigma \in G^{(j+3)} - S_m^{(j+2)}$. Se $y_p^{(q)}$ ocorreu no Passo $j-1$, então apenas substituímos essa variável pelo escalar $\lambda_p^{(q)}$ definido anteriormente. Caso contrário, substituímos $y_p^{(q)}$ por qualquer escalar $\lambda_p^{(q)}$ em $\mathbb{K}$. Dessa forma, obteremos uma função linear em termos de $y_{j+2}^{(j+2)}$:

$$\left(\cdots \left(\left(\sum_{\sigma \in G^{(4)}} \alpha_\sigma \lambda_2^{(\sigma(2))} \lambda_3^{(\sigma(3))} \right) \lambda_4^{(4)} + \sum_{\sigma G^{(5)} - S_m^{(4)}} \alpha_\sigma \lambda_2^{(\sigma(2))} \lambda_3^{(\sigma(3))} \lambda_4^{(\sigma(4))} \right) \lambda_5^{(5)} + \cdots \right) y_{j+2}^{(j+2)} \sum_{\sigma \in G^{(j+3)} - S_m^{(j+2)}} \alpha_\sigma \lambda_2^{(\sigma(2))} \cdots \lambda_{j+2}^{(\sigma(j+2))} \quad (5.10)$$

Pelo Passo $j-1$, o coeficiente de $y_{j+2}^{(j+2)}$ acima é não nulo. Portanto, podemos substituir $y_{j+2}^{(j+2)}$ por algum elemento $\lambda_{j+2}^{(j+2)}$ em $\mathbb{K}$ de sorte que (5.10) seja não nulo em $\mathbb{K}$.

Caso 2: Para $k = 2$ em (5.9) e para $\sigma \in G^{(j+2)}$, substituímos as variáveis $y_3^{(\sigma(2))}, \cdots, y_{j+2}^{(\sigma(j+1))}$ pelos escalares $\lambda_3^{(\sigma(2))}, \cdots, \lambda_{j+2}^{(\sigma(j+1))}$ definidos no Passo $j-1$, respectivamente. Agora consideramos todas as variáveis $y_p^{(q)}$ dentre $y_3^{(\sigma(2))}, \cdots, y_{j+3}^{(\sigma(j+2))}$, com $\sigma \in G^{(j+3)} - S_m^{(j+2)}$. Se $y_p^{(q)}$ ocorreu no Passo 1 ou no Caso 1, então apenas substituímos essa variável pelo escalar $\lambda_p^{(q)}$ definido anteriormente. Caso contrário, substituímos $y_p^{(q)}$ por qualquer escalar $\lambda_p^{(q)}$ em $\mathbb{K}$. Dessa forma, obteremos uma função linear em termos de $y_{j+3}^{(j+2)}$:

$$\left(\cdots\left(\left(\sum_{\sigma\in G^{(4)}} \alpha_\sigma \lambda_3^{(\sigma(2))}\lambda_4^{(\sigma(3))}\right)\lambda_5^{(4)} + \sum_{\sigma G^{(5)}-S_m^{(4)}} \alpha_\sigma \lambda_3^{(\sigma(2))}\lambda_4^{(\sigma(3))}\lambda_5^{(\sigma(4))}\right)\lambda_6^{(5)} + \cdots\right) y_{j+3}^{(j+2)}$$
$$\sum_{\sigma\in G^{(j+3)}-S_m^{(j+2)}} \alpha_\sigma \lambda_3^{(\sigma(2))}\cdots\lambda_{j+3}^{(\sigma(j+2))} \quad (5.11)$$

Pelo Passo $j-1$, o coeficiente de $y_{j+3}^{(j+2)}$ acima é não nulo. Portanto, podemos substituir $y_{j+3}^{(j+2)}$ por algum elemento $\lambda_{j+3}^{(j+2)}$ em $\mathbb{K}$ de sorte que (5.11) seja não nulo em $\mathbb{K}$.

Seguindo essas ideias, o último caso será:

Caso $n-m$: Para $k = n-m$ em (5.9) e para $\sigma \in G^{(j+2)}$, substituímos as variáveis $y_{n-m+1}^{(\sigma(2))}, \cdots, y_{n-m+j}^{(\sigma(j+1))}$ pelos escalares $\lambda_{n-m+1}^{(\sigma(2))}, \cdots, \lambda_{n-m+j}^{(\sigma(j+1))}$ definidos no Passo $j-1$, respectivamente. Agora consideramos todas as variáveis $y_p^{(q)}$ dentre $y_{n-m+1}^{(\sigma(2))}, \cdots, y_{n-m+j+1}^{(\sigma(j+2))}$, com $\sigma \in G^{(j+3)} - S_m^{(j+2)}$. Se $y_p^{(q)}$ ocorreu no Passo 1 ou nos casos anteriores, então apenas substituímos essa variável pelo escalar $\lambda_p^{(q)}$ definido anteriormente. Caso contrário, substituímos $y_p^{(q)}$ por qualquer escalar $\lambda_p^{(q)}$ em $\mathbb{K}$. Dessa forma, obteremos uma função linear em termos de $y_{n-m+j+1}^{(j+2)}$:

$$\left(\cdots\left(\left(\sum_{\sigma\in G^{(4)}} \alpha_\sigma \lambda_{n-m+1}^{(\sigma(2))}\lambda_{n-m+2}^{(\sigma(3))}\right)\lambda_{n-m+3}^{(4)} + \sum_{\sigma G^{(5)}-S_m^{(4)}} \alpha_\sigma \lambda_{n-m+1}^{(\sigma(2))}\lambda_{n-m+2}^{(\sigma(3))}\lambda_{n-m+3}^{(\sigma(4))}\right)\lambda_{n-m+4}^{(5)}\right.$$
$$\left. + \cdots\right) y_{n-m+j+1}^{(j+2)} + \sum_{\sigma\in G^{(j+3)}-S_m^{(j+2)}} \alpha_\sigma \lambda_{n-m+1}^{(\sigma(2))}\cdots\lambda_{n-m+j+1}^{(\sigma(j+2))} \quad (5.12)$$

Pelo Passo $j-1$, o coeficiente de $y_{n-m+j+1}^{(j+2)}$ acima é não nulo. Portanto, podemos substituir $y_{n-m+j+1}^{(j+2)}$ por algum elemento $\lambda_{n-m+j+1}$ em $\mathbb{K}$ de sorte que (5.12) seja não nulo em $\mathbb{K}$, encerrando assim a prova do lema. $\square$

5.2 A prova do resultado principal

Antes de expormos o enunciado e a prova do resultado principal deste capítulo, faremos a seguinte definição.

Definição 5.4. *Sejam $\mathbb{K}$ um corpo qualquer, $n \geqslant 2$ um inteiro e $i \in \{1, \ldots, n\}$. Diremos que uma matriz em UT_n é (i)-diagonal se as entradas $(l, l+(i-1))$ são as únicas possivelmente não nulas, em que $l \in \{1, \ldots, n-i+1\}$. Em outras palavras, uma matriz (i)-diagonal é aquela que possui a seguinte forma*

$$\begin{pmatrix} & & \lambda_{1,i} & & 0 \\ & & & \ddots & \\ & & & & \lambda_{n-i+1,n} \\ & & & & \\ 0 & & & & \end{pmatrix}.$$

Com uma sutil mudança na Definição 5.4, podemos também considerar matrizes (i)-diagonais com entradas em $\mathbb{K}[Y]$.

Observação 5.5. *Para $k \geqslant -1$ inteiro, toda matriz $A \in UT_n^{(k)}$ pode ser escrita como uma soma de matrizes (i)-diagonais, em que $i \in \{k+2, \ldots, n\}$. De fato, se $A = \sum_{p,q=1}^{n} a_{p,q} e_{p,q}$, então*

$$A = \sum_{i=k+2}^{n} A_i, \tag{5.13}$$

em que A_i é a matriz (i)-diagonal cuja entrada $(l, l+(i-1))$ é igual à $a_{l,l+(i-1)}$, para todo $l \in \{1, \ldots, n-i+1\}$.

Teorema 5.6. *Sejam $\mathbb{K}$ um corpo qualquer, $n \geqslant 2$ e $m \geqslant 2$ inteiros e $f(x_1, \ldots, x_m) \in \mathbb{K}\langle X\rangle$ um polinômio multilinear não nulo. Então, $f(UT_n^{(0)})$ é $\{0\}$ ou $UT_n^{(m-1)}$.*

Demonstração. Escreva $f(x_1, \ldots, x_m) = \sum_{\sigma \in S_m} \alpha_\sigma x_{\sigma(1)} \cdots x_{\sigma(m)}, \alpha_\sigma \in \mathbb{K}$. Sem perda de generalidade, podemos supor $\alpha_{id} = 1$.

Desde que $UT_n^{(0)}$ é uma álgebra nilpotente de índice n, então qualquer polinômio multilinear não nulo $f(x_1, \ldots, x_m)$ é uma identidade polinomial para $UT_n^{(0)}$, quando $m \geqslant n$. Observamos também que caso $n > m$, então $f(x_1, \ldots, x_n)$ não é uma identidade polinomial para $UT_n^{(0)}$, já que substituindo cada variável x_j por $e_{j,j+1}$ obtemos

$$f(e_{1,2}, \ldots, e_{m,m+1}) = e_{1,m+1} \neq 0.$$

Portanto, a partir de agora iremos também supor que $n > m$. Escrevemos,

$$f = \sum_{j=1}^{m} f_j, \tag{5.14}$$

em que cada f_j é a soma de todos os monômios de f cuja j-ésima variável é igual à x_1.

Tomando

$$x_1^{(m+1)} = \sum_{k=1}^{n-1} y_k^{(m+1)} e_{k,k+1}, x_2 = \sum_{k=1}^{n-1} y_k^{(2)} e_{k,k+1}, \ldots, x_m = \sum_{k=1}^{n-1} y_m^{(m)} e_{k,k+1} \tag{5.15}$$

como matrizes (2)-diagonais com entradas em $\mathbb{K}[Y]$, por (5.14) obtemos

$$f(x_1^{(m+1)}, x_2, \ldots, x_m) = \sum_{k=1}^{n-m} \Bigg(y_k^{(m+1)} \sum_{\sigma \in S_m^{(1)}} \alpha_\sigma y_{k+1}^{(\sigma(2))} \cdots y_{k+m-1}^{(\sigma(m))}$$

$$+ y_{k+1}^{(m+1)} \delta_2^{(m+1)}(x_2, \ldots, x_m) + \cdots + y_{k+m-1}^{(m+1)} \delta_m^{(m+1)}(x_2, \ldots, x_m) \Bigg) e_{k,k+m}, \tag{5.16}$$

em que $y_{k+j-1}^{(m+1)} \delta_j^{(m+1)}(x_2, \ldots, x_m)$ denota a entrada $(k, k+m)$ da matriz $f_j(x_1^{(m+1)}, x_2, \ldots, x_m)$, para $j \in \{2, \ldots, m\}$.

Agora, considerando $A_{m+1} = \sum_{k=1}^{n-m} a_k^{(m+1)} e_{k,k+m} \in UT_n^{(m-1)}$, procuremos por uma solução do seguinte sistema não linear:

$$\begin{cases} y_1^{(m+1)} \sum_{\sigma \in S_m^{(1)}} \alpha_\sigma y_2^{(\sigma(2))} \cdots y_m^{(\sigma(m))} + y_2^{(m+1)} \delta_2^{(m+1)}(x_2, \ldots, x_m) + \cdots \\ + y_m^{(m+1)} \delta_m^{(m+1)}(x_2, \ldots, x_m) = a_1^{(m+1)} \\ y_2^{(m+1)} \sum_{\sigma \in S_m^{(1)}} \alpha_\sigma y_3^{(\sigma(2))} \cdots y_{m+1}^{(\sigma(m))} + y_3^{(m+1)} \delta_2^{(m+1)}(x_2, \ldots, x_m) + \cdots \\ + y_{m+1}^{(m+1)} \delta_m^{(m+1)}(x_2, \ldots, x_m) = a_2^{(m+1)} \\ \vdots \\ y_{n-m}^{(m+1)} \sum_{\sigma \in S_m^{(1)}} \alpha_\sigma y_{n-m+1}^{(\sigma(2))} \cdots y_{n-1}^{(\sigma(m))} + y_{n-m+1}^{(m+1)} \delta_2^{(m+1)}(x_2, \ldots, x_m) + \cdots \\ + y_{n-1}^{(m+1)} \delta_m^{(m+1)}(x_2, \ldots, x_m) = a_{n-m}^{(m+1)} \end{cases} \tag{5.17}$$

Pelo Lema 5.1, podemos encontrar matrizes em $UT_n^{(0)}$

$$\bar{x}_2 = \sum_{l=1}^{n-1} \lambda_l^{(2)} e_{l,l+1}, \ldots, \bar{x}_m = \sum_{l=1}^{n-1} \lambda_l^{(m)} e_{l,l+1},$$

tais que $\sum_{\sigma \in S_m^{(1)}} \alpha_\sigma \lambda_{k+1}^{(\sigma(2))} \cdots \lambda_{k+m-1}^{(\sigma(m))}$ são não nulos, para todo $k \in \{1, \ldots, n-1\}$. Segue que o sistema (5.17) torna-se num sistema linear nas variáveis $y_k^{(m+1)}$, com $k \in \{1, \ldots, n-1\}$. Podemos resolver tal sistema recursivamente começando com a última equação: substituímos $y_{n-m+1}^{(m+1)}, y_{n-m+2}^{(m+1)}, \ldots, y_{n-1}^{(m+1)}$ por quaisquer valores em $\mathbb{K}$ (por exemplo por 0) e resolvemos essa equação para $y_{n-m}^{(m+1)}$. A seguir resolvemos a equação anterior para $y_{n-m-1}^{(m+1)}$, e continuamos com esse processo até obtermos uma solução do sistema. Portanto, qualquer matriz $(m+1)$-diagonal A_{m+1} pode ser realizada como $f(\bar{x}_1^{(m+1)}, \bar{x}_2, \ldots, \bar{x}_m)$, para alguma matriz $\bar{x}_1^{(m+1)}$ em $UT_n^{(0)}$.

Agora, para cada $i \in \{m+2, \ldots, n\}$, consideramos a matriz

$$x_1^{(i)} = \sum_{k=1}^{n-i+m} y_k^{(i)} e_{k,k+i-m}.$$

com entradas em $\mathbb{K}[Y]$. Então,

$$f(x_1^{(i)}, \bar{x}_2, \ldots, \bar{x}_m) = \sum_{k=1}^{n-i+1} \Bigg(y_k^{(i)} \sum_{\sigma \in S_m^{(1)}} \alpha_\sigma \lambda_{k+i-m}^{(\sigma(2))} \cdots \lambda_{k+i-2}^{(\sigma(2))}$$
$$+ y_{k+1}^{(i)} \delta_2^{(i)}(\bar{x}_2, \ldots, \bar{x}_m) + \cdots + y_{k+m-1} \delta_m^{(i)}(\bar{x}_2, \ldots, b\bar{x}_m) \Bigg) e_{k,k+i-1}$$

em que $y_{k+j-1}^{(i)} \delta_j^{(i)}(\bar{x}_2, \ldots, \bar{x}_m)$ denota a entrada $(k, m+k)$ da matriz $f_j(x_1^{(i)}, \bar{x}_2, \ldots, \bar{x}_m)$, para $j \in \{2, \ldots, m\}$.

Afirmamos que $\sum_{\sigma \in S_m^{(1)}} \alpha_\sigma \lambda_{k+i-m}^{(\sigma(2))} \cdots \lambda_{k+i-2}^{(\sigma(m))}$ é não nulo para $k \in \{1, \ldots, n-i+1\}$. De fato, podemos reescrever essa soma como $\sum_{\sigma \in S_m^{(1)}} \alpha_\sigma \lambda_{k+1}^{(\sigma(2))} \cdots \lambda_{k+m-1}^{(\sigma(m))}$, com $k \in \{i-m, \ldots, n-m\}$. Portanto, para cada $i \in \{m+2, \ldots, n\}$ e para qualquer matriz (i)-diagonal $A_i = \sum_{k=1}^{n-i+1} a_k^{(i)} e_{k,k+i-1} \in UT_n^{(m-1)}$ dada, uma solução do seguinte sistema

$$\begin{cases} y_1^{(i)} \sum_{\sigma \in S_m^{(1)}} \alpha_\sigma \lambda_{i+1-m}^{(\sigma(2))} \cdots \lambda_{i-1}^{(\sigma(m))} + y_2^{(i)} \delta_2^{(i)}(\bar{x}_2, \ldots, \bar{x}_m) + \cdots \\ + y_m^{(i)} \delta_m^{(i)}(\bar{x}_2, \ldots, \bar{x}_m) = a_1^{(i)} \\ y_2^{(i)} \sum_{\sigma \in S_m^{(1)}} \alpha_\sigma \lambda_{i+2-m}^{(\sigma(2))} \cdots \lambda_{i}^{(\sigma(m))} + y_3^{(i)} \delta_2^{(i)}(\bar{x}_2, \ldots, \bar{x}_m) + \cdots \\ + y_{m+1}^{(i)} \delta_m^{(i)}(\bar{x}_2, \ldots, \bar{x}_m) = a_2^{(i)} \\ \vdots \\ y_{n-i+1}^{(i)} \sum_{\sigma \in S_m^{(1)}} \alpha_\sigma \lambda_{n-m+1}^{(\sigma(2))} \cdots \lambda_{n-1}^{(\sigma(m))} + y_{n-i+2}^{(i)} \delta_2^{(i)}(\bar{x}_2, \ldots, \bar{x}_m) + \cdots \\ + y_{n-i+m}^{(i)} \delta_m^{(i)}(\bar{x}_2, \ldots, \bar{x}_m) = a_{n-i+1}^{(i)} \end{cases}$$

pode ser encontrada recursivamente.

Concluímos que qualquer matriz (i)-diagonal A_i pode ser realizada como $f(\bar{x}_1^{(i)}, \bar{x}_2, \ldots, \bar{x}_m)$ para alguma matriz $\bar{x}_1^{(i)}$ em $UT_n^{(0)}$.

Finalmente, dada uma matriz qualquer $A \in UT_n^{(m-1)}$, por (5.13) temos

$$A = \sum_{i=m+1}^{n} A_i = \sum_{i=m+1}^{n} f(\bar{x}_1^{(i)}, \bar{x}_2, \ldots, \bar{x}_m) = f(\sum_{i=m+1}^{n} \bar{x}_1^{(i)}, \bar{x}_2, \ldots, \bar{x}_m) \in f(UT_n^{(0)}).$$

Portanto, $UT_n^{(m-1)} \subset f(UT_n^{(0)})$. Sendo a inclusão contrária trivial, o teorema está provado. $\square$

A partir de agora iremos observar que analogamente à demonstração do Teorema 5.6 podemos obter um resultado mais geral, que está exposto no seguinte teorema.

Teorema 5.7. *Seja $f(x_1, \ldots, x_m) \in \mathbb{K}\langle X \rangle$ um polinômio multilinear não nulo. Sejam $n \geqslant 2$ um inteiro e $i \in \{0, \ldots, n-2\}$. Então, $f(UT_n^{(i)})$ é igual a $\{0\}$ ou $UT_n^{(m(i+1)-1)}$.*

Demonstração. Se $(m+i) \geqslant n$, então $f(UT_n^{(i)}) = \{0\}$. Logo podemos supor que $m(i+1) < n$ e assim iremos provar que $f(UT_n^{(i)}) = UT_n^{(m(i+1)-1)}$.

Também supomos $f(x_1, \ldots, x_m) = \sum_{\sigma \in S_m} \alpha_\sigma x_{\sigma(1)} \cdots x_{\sigma(m)}$ com $\lambda_{id} = 1$ e escrevemos f como em (5.14). Denotando $m(i+1) + 1 = l$ e tomando

$$x_1^{(l)} = \sum_{k=1}^{n-i-1} y_k^{(l)} e_{k,k+i+1}, x_2 = \sum_{k=1}^{n-i-1} y_k^{(2)} e_{k,k+i+1}, \ldots, x_m = \sum_{k=1}^{n-i-1} y_k^{(m)} e_{k,k+i+1}$$

temos

$$f(x_1^{(l)}, x_2, \ldots, x_m) = \sum_{k=1}^{n-m(i+1)} \Bigg(y_k^{(l)} \sum_{\sigma \in S_m^{(1)}} \alpha_\sigma y_{k+i+1}^{(\sigma(2))} \cdots y_{k+(m-1)(i+1)}^{(\sigma(m))}$$

$$y_{k+i+1}^{(l)} \delta_2^{(l)}(x_2, \ldots, x_m) + \cdots + y_{k+(m-1)(i+1)}^{(l)} \delta_m^{(l)}(x_2, \ldots, x_m) \Bigg) e_{k,k+m(i+1)},$$

em que $y_{k+(j-1)(i+1)}^{(l)} \delta_j^{(l)}(x_2, \ldots, x_m)$ denota a entrada $(k, k+m(i+1))$ da matriz

$$f_j(x_1^{(l)}, x_2, \ldots, x_m),$$

para $j = 2, \ldots, m$, e a soma $\sum_{\sigma \in S_m^{(1)}} \lambda_\sigma y_{k+i+1}^{(\sigma(2))} \cdots y_{k+(m-1)(i+1)}^{(\sigma(m))}$ foi obtida por $f_1(x_1^{(l)}, x_2, \ldots, x_m)$.

Considerando $B_l = \sum_{k=1}^{n-m(i+1)} b_k^{(l)} e_{k,k+m(i+1)} \in UT_n^{(m(i+1)-1)}$, procuramos por uma solução do seguinte sistema não linear

$$\begin{cases} y_1^{(l)} \sum\limits_{\sigma \in S_m^{(1)}} \alpha_\sigma y_{i+2}^{(\sigma(2))} \cdots y_{(m-1)i+m}^{(\sigma(m))} + y_{i+2}^{(l)} \delta_2^{(l)}(x_2, \ldots, x_m) + \cdots \\ + y_{(m-1)i+m}^{(l)} \delta_m^{(l)}(x_2, \ldots, x_m) = b_1^{(l)} \\ y_2^{(l)} \sum\limits_{\sigma \in S_m^{(1)}} \alpha_\sigma y_{i+3}^{(\sigma(2))} \cdots y_{(m-1)i+m+1}^{(\sigma(m))} + y_{i+3}^{(l)} \delta_2^{(l)}(x_2, \ldots, x_m) + \cdots \\ + y_{(m-1)i+m+1}^{(l)} \delta_m^{(l)}(x_2, \ldots, x_m) = b_2^{(l)} \\ \vdots \\ y_{n-m(i+1)}^{(l)} \sum\limits_{\sigma \in S_m^{(1)}} \alpha_\sigma y_{n-(m-1)(i+1)}^{(\sigma(2))} \cdots y_{n-(i+1)}^{(\sigma(m))} + y_{n-(m-1)(i+1)}^{(l)} \delta_2^{(l)}(x_2, \ldots, x_m) + \cdots \\ + y_{n-(i+1)}^{(l)} \delta_m^{(l)}(x_2, \ldots, x_m) = b_{n-m(i+1)}^{(l)} \end{cases}$$

Analogamente ao Lema 5.1, conseguimos obter matrizes em $UT_n^{(i)}$

$$\bar{x}_2 = \sum_{k=1}^{n-i-1} \lambda_k^{(2)} e_{k,k+i+1}, \ldots, \bar{x}_m = \sum_{k=1}^{n-i-1} \lambda_k^{(m)} e_{k,k+i+1},$$

tais que $\sum\limits_{\sigma \in S_m^{(1)}} \alpha_\sigma \lambda_{k+i+1}^{(\sigma(2))} \cdots \lambda_{k+(m-1)(i+1)}^{(\sigma(m))}$ são não nulos, para todo $k = 1, \ldots, n - m(i+1)$. Então podemos resolver o sistema linear obtido recursivamente.

Para cada $t \in \{m(i+1)+2, \ldots, n\}$, tomando $x_1^{(t)} = \sum\limits_{k=1}^{n-i-t+m(i+1)} y_k^{(t)} e_{k,k+i+t-m(i+1)}$, temos

$$f(x_1^{(t)}, \bar{x}_2, \ldots, \bar{x}_m) = \sum_{k=1}^{n-t+1} \Bigg(y_k^{(t)} \sum_{\sigma \in S_m^{(1)}} \alpha_\sigma \lambda_{k+i+t-m(i+1)}^{(\sigma(2))} \cdots \lambda_{k+t-i-2}^{(\sigma(m))}$$
$$+ y_{k+i+1}^{(t)} \delta_2^{(t)}(\bar{x}_2, \ldots, \bar{x}_m) + \cdots + y_{k+(m-1)(i+1)} \delta_m^{(t)}(\bar{x}_2, \ldots, \bar{x}_m) \Bigg) e_{k,k+t-1}$$

em que $y_{k+(j-1)(i+1)}^{(t)} \delta_j^{(t)}(\bar{x}_2, \ldots, \bar{x}_m)$ denota a entrada $(k, k+t-1)$ da matriz $f_j(x_1^{(t)}, \bar{x}_2, \ldots, \bar{x}_m)$, para $j = 2, \ldots, m$, e a soma $\sum\limits_{\sigma \in S_m^{(1)}} \lambda_\sigma \alpha_{k+i+t-m(i+1)}^{(\sigma(2))} \cdots \alpha_{k+t-i-2}^{(\sigma(m))}$ foi obtida por $f_1(x_1^{(t)}, \bar{x}_2, \ldots, \bar{x}_m)$. Além disso, para cada $t \in \{m(i+1)+2, \ldots, n\}$, afirmamos que $\sum\limits_{\sigma \in S_m^{(1)}} \alpha_\sigma \lambda_{k+i+t-m(i+1)}^{(\sigma(2))} \cdots \lambda_{k+t-i-2}^{(\sigma(m))}$ é não nulo para todo $k = 1, \ldots, n-t+1$. De fato, podemos reescrever essa soma como $\sum\limits_{\sigma \in S_m} \lambda_\sigma \alpha_{k+i+1}^{(\sigma(2))} \cdots \alpha_{k+(m-1)(i+1)}^{(\sigma(m))}$ com $k = t - m(i+1), \ldots, n - m(i+1)$.

Portanto, tomando $B_t = \sum_{k=1}^{n-t+1} b_k^{(t)} e_{k,k+t-1} \in UT_n^{(m(i+1)-1)}$, nós podemos considerar um sistema linear como em (5.18) e encontrar uma solução recursivamente. Desse modo, a conclusão da demonstração segue como a do Teorema 5.6. $\square$

REFERÊNCIAS

[1] A. Albert and B. Muckenhoupt. On matrices of trace zero. *Michigan Math. J.*, 4:1–3, 1957.

[2] S. Amitsur and L. Rowen. Elements of reduced trace 0. *Israel J. Math.*, 87:161–179, 1994.

[3] Z. Anzis, B. Emrich, and K. Valiveti. On the images of Lie polynomials evaluated on Lie algebras. *Linear Algebra Appl.*, 469:51–75, 2015.

[4] Y. Bahturin. *Identical relations in Lie algebras.* VNU Science Press b.v., Utrecht, 1987.

[5] M. Brešar. *Introduction to noncommutative algebra.* Springer, Switzerland, 2014.

[6] M. Brešar and I. Klep. Values of noncommutative polynomials, Lie skew ideals and tracial Nullstellensätze. *Math. Res. Lett.*, 16:605–626, 2009.

[7] D. Buzinski and R. Winstanley. On multilinear polynomials in four variables evaluated on matrices. *Linear Algebra Appl.*, 439:2712–2719, 2013.

[8] C. Chuang. On ranges of polynomials in finite matrix rings. *Proc. Amer. Math. Soc.*, 110:293–302, 1990.

[9] V. Drensky. *Free algebras and PI-algebras.* Springer, Singapore, 1999.

[10] K. Dykema and I. Klep. Instances of the Kaplansky Lvov multilinear conjecture for polynomials of degree three. *Linear Algebra Appl.*, 508:272–288, 2016.

[11] P. Fagundes. The images of multilinear polynomials on strictly upper triangular matrices. *Linear Algebra Appl.*, 563:287–301, 2019.

[12] P. Fagundes and T. de Mello. Images of multilinear polynomials of degree up to four on upper triangular matrices. *Oper. Matrices*, 2019, accepted.

[13] V. Filippov, V. Kharchenko, and I. Shestakov, editors. *The Dniester notebook: unsolved problems in the theory of rings and modules.* 4th ed. Mathematics Institute, Russian Academy of Sciences, Siberian Branch, Novosibirsk, 1993.

[14] M. Hall. A basis for free Lie rings and higher commutators in free groups. *Proc. Amer. Math. Soc.*, 1:575–581, 1950.

[15] I. Herstein. On Lie and Jordan rings of a simple associative ring. *Amer. J. of Math.*, 77:279–285, 1955.

[16] S. Kanel-Belov, A. Malev, and L. Rowen. The images of non-commutative polynomials evaluated on 2×2 matrices. *Proc. Amer. Math. Soc.*, 140:465–478, 2012.

[17] S. Kanel-Belov, A. Malev, and L. Rowen. The images of multilinear polynomials evaluated on 3×3 matrices. *Proc. Amer. Math. Soc.*, 144:7–19, 2016.

[18] S. Kanel-Belov, A. Malev, and L. Rowen. The images of Lie polynomials evaluated on matrices. *Commun. Algebra*, 45:4801–4808, 2017.

[19] A. Kostrikin and Y. Manin. *Linear Algebra and Geometry.* Algebra, Logic and Applications, CRC Press, Translated from Russian by Alferieff, Moscow, 1989.

[20] A. Ma and J. Oliva. On the images of Jordan polynomials evaluated over symmetric matrices. *Linear Algebra Appl.*, 492:13–25, 2016.

[21] A. Malev. The image of non-commutative polynomials evaluated on 2×2 matrices over an arbitrary field. *J Algebra Appl.*, 13, 1450004:12 pp., 2014.

[22] Z. Mesyan. Polynomial of small degree evaluated on matrices. *Linear Multilin. Alg.*, 61:1487–1495, 2013.

[23] M. Rosset and S. Rosset. Elements of trace zero that are not commutators. *Commun. Algebra*, 28:3059–3072, 2000.

[24] K. Shoda. Einige sätze über matrizen. *Jap. J. Math.*, 13:361–365, 1936.

[25] Š. Špenko. On the images of a noncommutative polynomial. *J. Algebra*, 377:298–311, 2013.